Methodological Issues In Oral Health Related Quality Of Life Research Using The Oral Health Impact Profile (OHIP-14)

Andiappan, Manoharan

Awarding institution:
King's College London

Methodological Issues In Oral Health Related Quality Of Life Research Using The Oral Health Impact Profile (OHIP-14)

by

Manoharan Andiappan

Thesis submitted for the Degree of

Doctor of Philosophy

Dental Institute

King's College London

2018

Acknowledgements

At the outset, I would like to express my deep sense of gratitude to my Supervisors Prof. Francis Hughes and Dr. Wei Gao for their continued support and guidance in carrying out this research.

I sincerely thank my previous supervisors Prof. Ana Nora Donaldson, without whom this project would not have materialised and Prof. Stephen Dunne for their support and guidance during the initial stages of this study and their encouragement.

My thanks are due to Dr. Helena Lewis Greene for her help and support during the data collection. My sincere thanks are due to Prof. Tim Newton and Dr. Adam Hasan for their critical comments and suggestions to improve the quality of this work.

I am grateful to Prof. Jennifer Gallagher and staff of Population and Patient Health Division for their support while carrying out this research. I acknowledge the help of my friend Dr. Ghotane Swapnil during the various stages of this research.

I thank my wife, children and other family members for their encouragement, understanding and continued support during the course of this study.

Thanks and thoughts of healing are also due to all the patients who kindly consented to take part in this study and provided data.

Dedicated to my beloved parents

TABLE OF CONTENTS

1 INTRODUCTION AND LITERATURE REVIEW 16
1.1 INTRODUCTION 16
1.2 QUALITY OF LIFE (QOL) 18
1.2.1 What is Quality of Life? 18
1.2.2 Defining Oral Health Related Quality of Life 19
1.2.3 Importance of Oral Health Related Quality of Life 20
1.3 CONCEPTUALISATION OF ORAL HEALTH RELATED QUALITY OF LIFE........... 21
1.3.1 Models of Health/ Oral Health Related Quality of Life.........................22
1.3.2 Quantification of Health Related Quality of Life and Oral Health Related Quality of Life 26
1.3.3 Measures for OHRQoL 29
1.3.4 Relationship between OHRQoL and demographic factors....................35
1.4 ORAL HEALTH IMPACT PROFILE -14 (OHIP-14) 38
1.5 METHODOLOGICAL CHALLENGES OF QUESTIONNAIRE BASED DATA 43
1.5.1 Missing data 43
1.5.2 Floor and Ceiling effects...44
1.5.3 Factorial structure 47
1.5.4 Responsiveness 47
1.6 METHODOLOGICAL ISSUES IN OHIP-14 DATA... 51
1.6.1 Missing Data 51
1.6.2 Floor and Ceiling effects in OHIP-14 data 55
1.6.3 Factorial structure of OHIP items 56
1.6.4 Responsiveness to change...60
1.6.5 Chapter Summary 63
1.7 AIMS AND OBJECTIVES.. 64

2 MATERIALS AND METHODS 66
2.1 CHAPTER INTRODUCTION 66
2.2 STUDY POPULATION.. 66
2.3 INCLUSION AND EXCLUSION CRITERIA 66
2.4 SAMPLE SIZE CONSIDERATION 67
2.5 ETHICAL APPROVAL 72
2.6 DATA COLLECTION... 72

2.6.1 Method of Data collection......72
2.6.2 Baseline Data......73
2.6.3 Data Collection Instruments......74
2.6.4 Follow-up-1......77
2.6.5 Follow-up-2......78

2.7 DATA HANDLING......79

2.7.1 Data checking and cleaning......79
2.7.2 Data coding and computerisation......79

2.8 STATISTICAL ANALYSIS......82

3 DESCRIPTION OF SAMPLE AND OHIP SCORES......86

3.1 CHAPTER INTRODUCTION......86

3.2 DESCRIPTIVE AND DEMOGRAPHIC DETAILS......86

3.3 DESCRIPTION OF ORAL HEALTH RELATED QUALITY OF LIFE FOR COMPLETED CASES......89

3.3.1 Predictors of Oral Health Related Quality of Life based on completed cases......90
3.3.2 OHRQoL by seven dimensions......92

3.4 PSYCHOMETRIC PROPERTIES OF OHIP-14......95

3.5 DISCUSSION AND CONCLUSIONS......98

4 MISSING DATA ANALYSIS......101

4.1 INTRODUCTION......101

4.2 METHODS FOR DEALING WITH MISSING DATA......104

4.2.1 Listwise Deletion......104
4.2.2 Mean substitution......104
4.2.3 Interpolation, Trend and Regression methods......105
4.2.4 Expectation Maximization Algorithm (EM Algorithm)......106
4.2.5 Multiple Imputation (MI)......106

4.3 STUDY METHODS......108

4.3.1 Missing data at baseline......108
4.3.2 Missing Data Pattern in the Follow-up data......109

4.4 RESULTS......111

4.4.1 Exploring missing OHIP items in baseline data......111
4.4.2 Comparison of Methods......115
4.4.3 Selection of suitable method for handling missing data in OHIP-14 items......117
4.4.4 Missing data in the follow-up study......118

4.4.5 Exploring dropout mechanisms 120
4.5 DISCUSSION.. 126

5 **FLOOR AND CEILING EFFECTS OF THE OHIP-14 ....130**

5.1 CHAPTER INTRODUCTION 130

5.2 METHODS 131

5.3 FLOOR AND CEILING EFFECTS 134

5.4 MODEL COMPARISONS 139

5.5 DISCUSSION.. 141

6 **FACTORIAL STRUCTURE OF THE OHIP-14 INSTRUMENT 145**

6.1 CHAPTER INTRODUCTION 145

6.2 METHODS 146

6.3 RESULTS 151

6.4 DISCUSSION.. 159

7 **RESPONSIVENESS OF OHIP-14..162**

7.1 CHAPTER INTRODUCTION.. 162

7.2 METHODS 162

7.2.1 Conventional method 162
7.2.2 Reliability of change scores 163
7.2.3 Reliable Change Index (RCI)..164
7.2.4 Simulated data for comparison 167
7.3 RESULTS 168

7.3.1 Descriptive Statistics...168
7.3.2 Aggregate change..169
7.3.3 Individual Change scores 171
7.4 DISCUSSION.. 174

8 **VALIDATION OF METHODOLOGY WITH A LARGE INDEPENDENT DATASET 177**

8.1 CHAPTER INTRODUCTION 177

8.2 DESCRIPTION OF NATIONAL SAMPLE 177

8.3 VALIDATION OF METHODOLOGY 178

8.3.1 Missing Data in Reference data set 178
8.3.2 Floor and Ceiling Effects 186
8.3.3 Factorial Structure 194

8.4 RESULTS 194

8.5 DISCUSSION 202

9 GENERAL DISCUSSION 205

9.1 OVERVIEW 205

9.1.1 Missing data 206
9.1.2 Floor and Ceiling Effects 211
9.1.3 Change scores 214
9.1.4 Factorial structure 218
9.1.5 Comparison of the findings from the current study and National sample 221

9.2 STRENGTHS AND WEAKNESSES OF THIS STUDY 222

9.3 FUTURE WORK 224

9.4 SUMMARY AND CONCLUSIONS 226

REFERENCES ……………………………………………..…………...230

APPENDICES…….…………………………………………...…………...249

TABLE OF FIGURES

Figure 1.1 Locker's conceptual model of oral health linking disease with the domains of oral health. 23

Figure 1.2 Wilson and Cleary conceptual model postulating that environmental and individual factors influence subjective OHQoL. 24

Figure 1.3 Gilbert's Conceptual Model of the multidimensionality of Oral Health and OHRQoL 25

Figure 1.4 Components involved in measuring Oral Health Related Quality of Life 31

Figure 3.1 Box plot depicting the median OHIP-14 total score for different ethnic groups. 89

Figure 3.2 The mean and 95% Confidence interval of total OHIP scores for various treatment needs 94

Figure 3.3 The mean and 95% Confidence interval of total OHIP scores for various Professions. 95

Figure 4.1 Processes involved in Multiple Imputation 107

Figure 4.2 Missing variables, cases and values in OHIP items 113

Figure 4.3 Missing data patterns in the data 115

Figure 4.4 Mean and 95% CI for dropouts and non-dropouts. 121

Figure 4.5 Predicted mean and 95% confidence interval at three time points for different treatment groups. 126

Figure 5.1 Total number of samples who scored lowest (0) (Floor) and highest (5) (ceiling) Scores in each of OHIP-14 items. 134

Figure 6.1 Fitted Single factor Structural Equation Model for OHIP-14 items. depicting the relationship between exogenous variables (factors) and endogenous variables (OHIP items) along with covariances. 153

Figure 6.2 Three factor model with estimates 154

Figure 6.3 Six factor model with estimates 155

Figure 6.4 Seven factor model with estimates 156

Figure 8.1 Bland Altman graphs for comparing the two measurements (Original and imputed total OHIP scores) for different missing data handling techniques. 184

Figure 8.2 Bland Altman graphs for comparing the two measurements (Original and imputed) for different missing data handling techniques. 185

Figure 8.3 The bar represents the number of participants who have given lowest (0) (the green bar) and highest (5) scores (the red bar) for each of OHIP-14 items...187

Figure 8.4 Fitted Single factor Structural Equation Model for OHIP-14 items. depicting the relationship between exogenous variables (factors) and endogenous variables (OHIP items) along with covariances. 196

Figure 8.5 Three factor model with estimates 197

Figure 8.6 Six factor model with estimates 198

Figure 8.7 Seven factor model with estimates 199

LIST OF TABLES

Table 1.1 Strengths and weaknesses of generic and specific measures. 27

Table 1.2 List of Generic measures and their dimensions 28

Table 1.3 List of Specific measures and their dimensions. 29

Table 1.4 Oral Health Related Quality of Life Questionnaires................................33

Table 1.5 Selected list of publications showing the relationship between OHIP-14 scores and demographic variables. 37

Table 3.1 Demographic details of the study sample 88

Table 3.2 Results of multivariate regression for predicting total OHIP score 91

Table 3.3 Dimension-wise summary of OHIP scores..93

Table 3.4 Bivariate correlation between different OHIP-14 items 96

Table 3.5 Item total statistics showing the relationship and the impact of each item in OHIP-14 instrument. 97

Table 4.1 Item wise percentage of missing data. 112

Table 4.2 Results of Logistic Regression to find out the predictors of missingness 114

Table 4.3 Estimates using different missing data techniques 116

Table 4.4 Mean and standard deviation of OHIP scores at various follow-up points by demographic details. 120

Table 4.5 Results of Logistic regression model to find out whether subject characteristics predict the dropouts. 122

Table 4.6 Results of individual Random Effects Models to find out the significant predictor 123

Table 4.7 Results of Random Effects Models 124

Table 5.1 Results of OLS and Tobit model fitting. 136

Table 5.2 Model fit statistics for OLS and Tobit models. 140

Table 6.1 Allocation of items to different factors in the 3 Factor Model (Montero et al 2010) 148

Table 6.2 Allocation of items to different factors in the 6 Factor Model 149

Table 6.3 Allocation of items to different factors in the 7 Factor Model 150

Table 6.4 Standardised Regression weights for different models. 157

Table 6.5 Variance estimates for different models 158

Table 6.6 Fit Indices for various models 159

Table 7.1 Mean and SD of composite scores at three time points. 168

Table 7.2 Change in OHIP score between Follow-up points................................170

Table 7.3 Classification of patients using Responsiveness measures 171

Table 7.4 Summary statistics of baseline scores for the three groups classified using RCI and SEoM based on the change scores from baseline to follow-up-2...173

Table 8.1 Estimates of Mean and sd values using different missing data techniques (Based on ADHS 2009 sample) 180

Table 8.2 Summary statistics of difference between original and imputed values182

Table 8.3 Bland - Altman statistics for imputed data from different methods 183

Table 8.4 Results of OLS and Tobit models using ADHS 2009 data 188

Table 8.5 Model fit statistics for OLS and Tobit models (For comparison data set) 193

Table 8.6 Standardised Regression weights for different models. 200

Table 8.7 Variance estimates for different models 201

Table 8.8 Fit Indices for four different models. 202

Abbreviations

ADHS	Adult Dental Health Survey
AMOS	Analysis of Moments Structure
CFI	Comparative Fit Index
COHQoL	Child Oral Health Quality of Life
CPITN	Community Periodontal Index of Treatment Needs
EM	Expectation Maximisation
EQ-5D+	Euro Quality of Life on 5 Dimensions
ES	Effect Size
EuroQOL	European Quality of Life
GOHAI	Geriatric Oral Health Assessment Index
HRQoL	Health Related Quality of Life
MAR	Missing At Random
MCAR	Missing Completely At Random
MI	Multiple Imputation
MID	Minimal Important Difference
MLE	Maximum Likelihood Estimate
MNAR	Missing Not At Random
NOHSA 2000	National Oral Health Survey of Adults 2000
OHIP	Oral Health Impact Profile
OHIP-14	OHIP with 14 items
OHIP-49	OHIP with 49 items
OH-QoL	Oral Health – Quality of Life
OHRQoL	Oral Health Related Quality of Life

OIDP	Oral Impact on Daily Performance
OLS	Ordinary Least Squares
PCA	Principal Component Analysis
PDC	Primary Dental Care
PIS	Patient Information Sheet
PRO	Patient Reported Outcome
QoL	Quality of Life
RC	Reliable Change
RCI	Reliable Change Index
RCSC	Reliable and Clinically Significant Change
Sd	Standard Deviation
SEdiff	Standard Error of difference
SEM	Structural Equation Modelling
SEoM	Standard Error of Measurement
SRS	Simple Random Sample
TMD	Temporo Mandibular Disorders
UKDA	United Kingdom Data Archive
VAS	Visual Analogue Scale
WHO	World Health Organisation

Abstract

This study aims to identify and address the methodological issues that may occur in the analysis of data for assessing Oral Health Related Quality of Life using the OHIP-14 instrument. The four main methodological issues addressed in this work are the handling of missing data, presence and management of floor and ceiling effects, number of dimensions in OHIP-14 and the responsiveness of OHIP items to change.

A total of 360 participants who came for dental treatment at King's College London Dental Hospital, Denmark Hill, London participated in this study. Baseline data were collected from participants at the time of treatment. Data were also collected at two follow-ups, two and four months after baseline. At baseline, data were collected from all the 360 participants whereas in the first and second follow-ups, 89 and 75 patients respectively provided data. Different techniques for managing missing data, namely completed case, Item mean, subject mean, interpolation, regression, trend, EM algorithm and multiple imputation were tested. The floor and ceiling effects were handled using the Tobit model. Structural Equation Modelling was used to test the existence of one, three, six and seven factor models and these models were compared.

The missing data in OHIP items followed a missing completely at random (MCAR) pattern. The mean values obtained from different missing data handling techniques were similar. No significant difference in mean OHIP scores was observed between dropout and non dropout cases and the dropouts

followed a Missing At Random (MAR) pattern. Education, Profession and treatment needs significantly predicted ($p<0.05$) the change in OHIP scores.

There was a greater floor effect than the ceiling effect. Use of the Tobit model, to adjust for floor and ceiling effects showed improved estimates for the effect of predictors. The comparison of Ordinary Least Squares (OLS) and Tobit model revealed that the Tobit model fitted the data well. OHIP-14 has good psychometric properties with the Cronbach's alpha value of 0.93 for measuring the OHRQoL. None of the four models identified from the literature (one, three, six and seven factor models) fitted the data well. OHIP-14 was responsive to change and the individuals were classified as "Improved", "No Change" and "Worsened" groups. The results were tested with national data from the Adult Dental Health Survey 2009, UK which showed similar results.

In conclusion, the missing data in OHIP items can be handled either by multiple imputation or EM algorithm and OHIP-14 items suffer from floor and ceiling effects which can be handled with the Tobit model. As none of the four models reported in the literature fitted the data well, further research is required to explore the dimensions of OHIP-14. OHIP-14 is responsive to change and can be used to measure the treatment effect over a period of time.

1 Introduction and Literature Review

1.1 Introduction

In this chapter the concept of Quality of Life is reviewed and in particular Health Related Quality of Life (HRQoL). More specifically Oral Health Related Quality of Life is explored in detail. The challenges of assessing OHRQoL, including the methodological problems that researchers may encounter while analysing the questionnaire based data will be discussed. Based on this review, a list of methodological problems is identified to address in this research.

The concept of Quality of Life (QoL) dates back to the early 19th century and was initially applied to patients with neoplastic disease (Leplege and Hunt, 1997). Quality of life has been viewed from three perspectives as i) the quality of one's life conditions ii) one's satisfaction with life conditions and iii) the combination of both (Borthwick-Duffy, 1992).

The Quality of Life (QoL) of an individual depends on the combination of life conditions and satisfaction but more emphasis to be given on personal values, aspirations and expectations (Felce and Perry, 1995). As health can play a major role in deciding an individual's life, its contribution towards quality of life is important. Initially, health was seen as a state of normal function that could be disrupted from time to time by disease. But its importance in an individual's life was subsequently realised. The specific impact of health on quality of life is termed Health Related Quality of Life (HRQoL). In the literature both Quality of

Life and Health Related Quality of Life are often used interchangeably. However, QoL is a broader concept and includes all aspects of life whereas HRQoL focus on the role of health, illness and the impact of medical treatment on QoL (Guyatt et al., 2007). Not surprisingly, given the importance of assessing QoL in determining both disease impacts and health outcomes, methods of objectively assessing or measuring QoL have been tested for a number of years. The subjective nature of health makes it difficult to measure directly and hence an indirect way to assess this by asking carefully a variety of structured questions has been initiated.

Measuring HRQoL is important as it helps to determine the burden of preventable disease, injuries, and disabilities, and can provide valuable new insights into the relationships between HRQOL and risk factors. It also helps to monitor progress in achieving health policy objectives (National Center for Chronic Disease Prevention and Health Promotion, USA). There has, in recent years, been increased focus and continuing research into the assessment of quality of life due to continual change in peoples' perception, change in people's settings and the need to measure the effectiveness of treatment.

Oral health is a subset of general health and oral problems may relate to the general health of an individual and can influence quality of life. Within the oral and dental field there has been considerable attention paid to the concept of Oral Health Related Quality of Life (OHRQoL). The American Dental Association define oral health as a "functional, structural, aesthetic, physiologic and psychosocial state of well-being and is essential to an individual’s general health and quality of life” (American Dental Association (ADA), 2014). Measuring Oral Health Related Quality of Life in the process of assessing treatment effectiveness and treatment planning has received wide attention in recent years and poses a number of challenges including *inter alia* reliability and validity of these assessments together with issues surrounding the analysis of data from such studies. Therefore, the focus of the present study is on the OHRQoL and some of the methodological issues associated with its measurement.

1.2 Quality of Life (QoL)

1.2.1 What is Quality of Life?

According to the definition given by World Health Organisation, quality of life is “*an individual’s perception of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards and concerns. It is a broad ranging concept affected in a complex way by the person’s physical health, psychological state, personal beliefs, social relationships and their relationship to salient features of their environment* ” (Oort, 2005). Hence, quality of life is a broad topic that includes various aspects of humans such as culture, psychology, environment etc. Quality of life is the gap between expectations and experience (Calman, 1984). This means that expectations differ according to individuals and hence the experiences differ. A person with a complicated medical condition may have a better quality of life than the one with a mild medical condition. Therefore, the quality of life perceived by individuals can vary over time and between people (Carr et al., 2001).

Health Related Quality of Life

Since the 1980s there has been increasing emphasis on the investigation of health-related quality of life (HRQOL) and its determinants and its impact on overall quality of life (Gandek et al., 2004, McHorney, 1999, Selim et al., 2009). Health-related quality of life also depends on many factors such as opportunity/resilience (favourable condition or the ability to recover from illness), health perceptions (how an individual perceives their health), functional states, impairments/diseases, and duration of life. The health-related quality of life approach has provided the opportunity for investigation of the interrelations among oral health, health, and related outcomes (Gift and Atchison, 1995). A significant association between HRQoL and Oral Health status has been reported (Naito et al., 2006) and hence oral health status can affect the overall health status of an individual.

1.2.2 Defining Oral Health Related Quality of Life

The consequences of oral diseases are not only physical but also economic, social and psychological and seriously impair the quality of life and affect various aspects of life (Naito et al., 2006). Therefore, the need to consider oral health as an integral part of health has been emphasised and the contribution of oral health towards the overall HRQoL recognised (Gift and Atchison, 1995). The World Health Organisation (WHO) has recognised oral health as an important part of the Global Health Programme as Oral Health Related Quality of Life (OHRQoL) is an integral part of general health (Petersen, 2003). Oral health problems have social, economic and psychological consequences, which mean that they have impact on the quality of life (Virdi, 2015). Oral health can affect people physically and psychologically and can influence many aspects as how they enjoy life, speak, chew, taste food, socialize and their social well-being (Locker, 1997), The evaluation of OHRQoL "reflects people's comfort when eating, sleeping and engaging in social interaction; their self-esteem; and their satisfaction with respect to their oral health" (Scully, 2000). OHRQoL is the way oral health affects the quality of life of an individual related to oral function, psychological well-being, social well-being and pain/discomfort (Inglehart and Bagramian, 2002). OHRQoL is defined as "*the impact of oral disease and disorders on aspects of everyday life that a patient or person values, that are of sufficient magnitude, in terms of frequency, severity or duration to affect their experience and perception of their life overall* " (Locker and Allen, 2007a, p.409).

1.2.3 Importance of Oral Health Related Quality of Life

Oral disease is an important public health problem due to the high prevalence of oral disease and the fact that it has significant impacts on the quality of life of an individual. Assessment of OHRQoL helps to understand and shape not only the state of clinical practice, dental research and dental education but also that of the community at large (Bennadi and Reddy, 2013). It has been reported that patient-reported outcomes help to understand the relationship between oral health and general health (Sischo and Broder, 2011). It has also been concluded that improving the quality of patients' well being may require more than simply treating dental problems alone. Quality of life issues are at the forefront in public health policy (Slade, 2002). The impact of commonly occurring oral disorders on people's functional, psychological and social well being has been reported by Slade (Slade, 2012). Thus the benefits of measuring OHRQoL have been reported by Gift and co-workers (Gift and Atchison, 1995) as:

1. to clinical practitioners in selecting treatments and monitoring patient outcomes;
2. to researchers in identifying determinants of health, tracking levels of health risk factors, and determining use of services in populations;
3. to policy makers in establishing programme and institutional priorities, policies and funding decisions.

1.3 Conceptualisation of Oral Health Related Quality of Life

The conceptualisation of OHRQoL using various theoretical models is reviewed in this section. As this study is based on data collected from adults, the current review is mainly focussed OHRQoL on adults and excludes studies that used children of less than 18 years old.

As people started to realise the importance of oral health, attempts to conceptualise oral health have also increased. Early assessments of oral health were mainly based on the actual conditions and did not depend on the respondent's perceptions, that is, how the patient feels about their health condition rather than simply the presence or absence of that condition. In addition, dentists had been trained in such measures of dental disease only. For example various indices have been used like Helkimo's index of mandibular dysfunction (Helkimo, 1974) and the Community Periodontal Index of Treatment Needs (CPITN) (Ainamo et al., 1982). These measures were used to find out the prevalence of the disease in the population. However the inability of these measures to reflect a patient's perception of disease (Gooch et al., 1989, Locker, 1988, Locker and Miller, 1994) indicated the necessity to develop a paradigm comprising different aspects of health (Allen, 2003). Hence, it has been widely accepted that any measure of health must include the social and emotional well-being of an individual along with the presence and absence of the disease. Based on these concepts, various conceptual frameworks have evolved. Three such important conceptual models are reviewed in chronological order below.

1.3.1 Models of Health/ Oral Health Related Quality of Life

Based on the WHO (1980) classification of impairment, disability and handicap, Locker (Locker, 1988) proposed a conceptual model for measuring oral health status of an individual. This model tries to record the psychosocial and functional consequences of illness where sickness, disability, functional limitation, and social disadvantage are linearly connected but can be modified by heterogeneous psychological and social conditions (Piovesan et al., 2009). According to this model, if an individual loses a tooth then he is impaired and hence he/she experiences discomfort in performing their routine and has functional limitation. This model provided the basis for further extensive work on OHRQoL by many researchers. Locker's model is depicted in Figure 1.1.

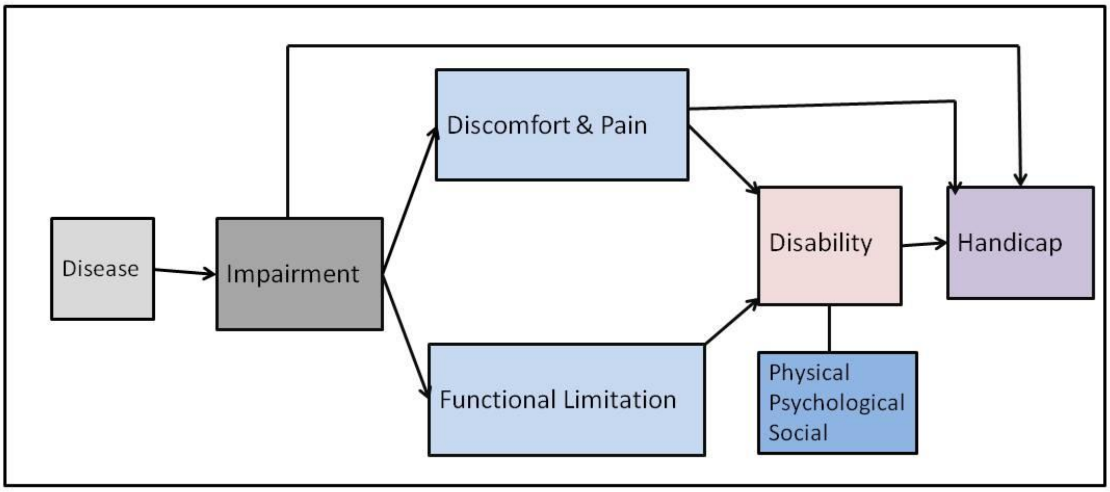

Figure 1.1 Locker's conceptual model of oral health linking disease with the domains of oral health.

The conceptual model developed by Wilson and Cleary (Wilson and Cleary, 1995) for patient outcomes has also been widely used in assessing the OHRQoL (Baker et al., 2008a, Benson et al., 2014, Gupta et al., 2015). This multi-dimensional model relates the clinical status to individual experience. In this model, they postulate that subjective oral health outcomes such as OHRQoL are influenced by environmental and individual factors. It encompasses disease, health and quality of life and links the causal relationship between them and provides a theoretical bridge between biomedical and socio-environmental concepts of health and modes of thinking. Based on this model and the model proposed by Brunner and Marmot (Brunner and Marmot, 2005), Gupta and co-workers (Gupta et al., 2015) have demonstrated that higher Socio-Economic Status predicted better OHRQoL mediated through a higher sense of coherence, higher social support and lower stress. The Wilson and Cleary model is depicted in Figure 1.2.

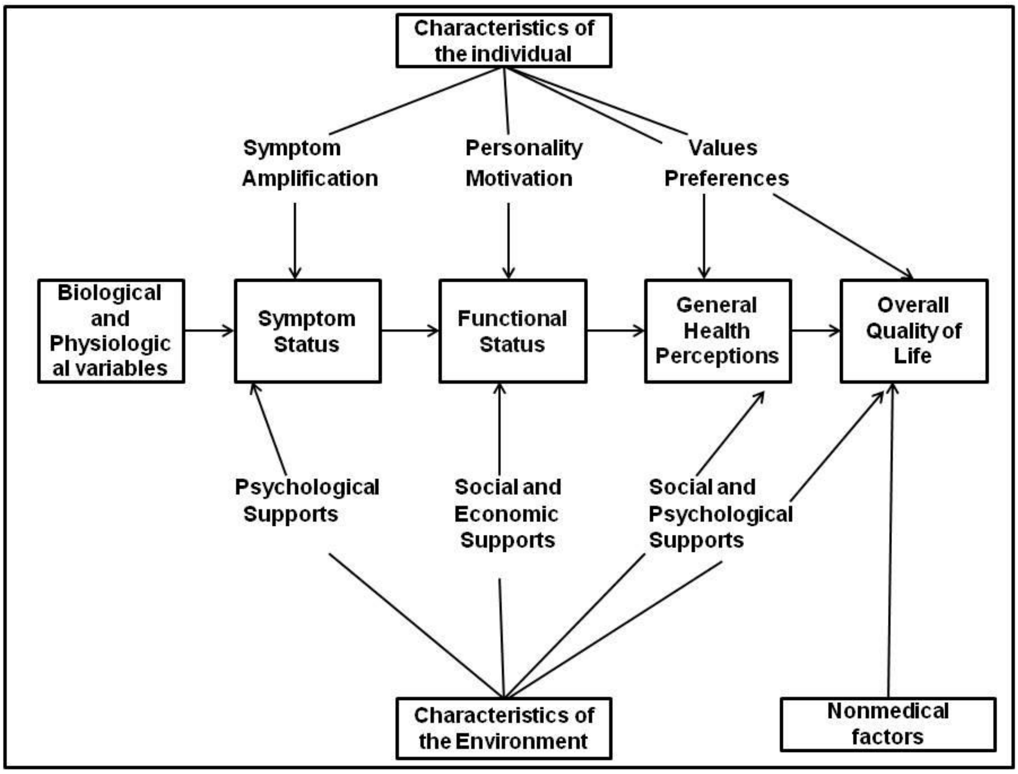

Figure 1.2 Wilson and Cleary conceptual model postulating that environmental and individual factors influence subjective OHQoL.

[Source: JAMA 1995; 273: 59–65. Copyright 1995, American Medical Association.]

As Oral Health Related Quality of Life is multidimensional, the conceptual framework incorporating the multi-dimensionality of Oral Health and Oral Health

Related Quality of Life was proposed (Gilbert, 2005). The five dimensional model proposed by Gilbert is as below in Figure 1.3.

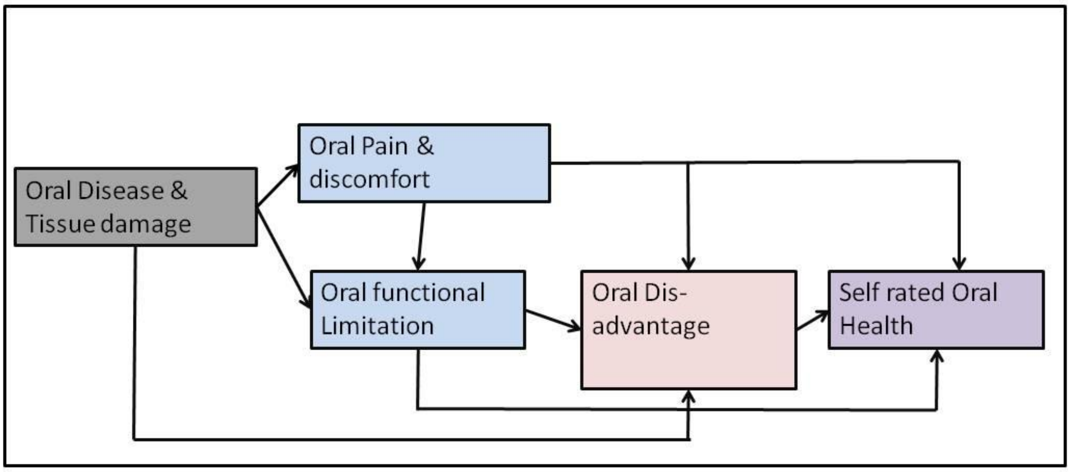

Figure 1.3 Gilbert's Conceptual Model of the multidimensionality of Oral Health and OHRQoL

The five dimensions of oral health and OHRQoL of Gilbert's model are:

- Oral disease and tissue damage,
- Oral pain and discomfort,
- Oral functional limitation,
- Oral disadvantage and
- Self-rated oral health.

The disease and tissue damage are measured by clinical examination and self-report. Gilbert has concluded that this model can be a base for further research in social disparities and for developing policy initiatives in order to eliminate the social disparities. Astrom and co-workers (Åstrøm et al., 2010) tested Gilbert's conceptual model (Gilbert, 2005, Gilbert et al., 1998a, Gilbert et al., 1998b) with three domains, namely symptom status, functional status and oral disadvantages and concluded that this model fitted their data well.

1.3.2 Quantification of Health Related Quality of Life and Oral Health Related Quality of Life

Quantification of HRQoL

Methods of quantification have received much attention because the measurement of quality of life has been identified as a significant indicator of service need and intervention outcome (Ng and Leung, 2006). In the process of quantification, a series of methodological issues need to be addressed before arriving at a standard measure. Health related quality of life instruments deal with generic and specific measures. The generic measures provide a summary measure of health while the specific measures focus on a particular disease condition and measure the patients' perceptions on health needs or to measure the outcome of interventions (Cunningham et al., 2001). Disease-specific measures assess the special states and concerns of diagnostic groups. Specific measures may be more sensitive for the detection and quantification of small changes that are important to clinicians or patients (Patrick and Deyo, 1989). Both measures have their own strengths and weaknesses which are summarised below in Table 1.1

Table 1.1 Strengths and weaknesses of generic and specific measures.

	Strength	Weakness
Generic	• Single instrument	• May not focus adequately on area of interest.
	• Comparison across different interventions or conditions is possible	• May not be sufficiently responsive
	• May be useful when condition specific measures are not available	• Some questions will be irrelevant.
	• Detect differential effects on different aspects of health status.	• Some generic instruments are excessively long.
Specific	• Clinically sensible	• Does not allow cross condition comparisons.
	• More responsive	• May be limited in terms of Populations and interventions.
	• More acceptable to patients, as they cover only relevant areas.	• Developmental process (reliability, validity and testing are time consuming.
	• Usually shorter than generic measures.	• More expensive to develop, administer and score.

[Source: (Bennett and Phillips, 1998, Guyatt et al., 1993)]

There are many instruments that have been developed for assessment of both generic and specific measures. Generic measures may provide operational definitions of several concepts summarized by a single index value or in a profile of interrelated scores (Patrick and Deyo, 1989). Generic measures provide summary of health-related quality of life and sometimes generate a single index measure of health (Cunningham et al., 2000). A list of generic measures is summarised in the Table 1.2.

Table 1.2 List of Generic measures and their dimensions

SNo	Measure	Dimensions measuring
1	Quality of Well being Scale (QWB) (Buschke et al., 1975)	Self-care, mobility, institutionalization, social activities, reports of symptoms and problems, including mental.
2	General Health Rating Index (Ware et al., 1978)	Six dimensions: past, present, and future 29 items perceptions of health; health-related worry and concern; resistance vs susceptibility to illness; tendency to view illness as a part of life.
3	Sickness Impact Profile (Bergner et al., 1981)	Physical: ambulation, mobility, body care , Psychosocial: social interaction, communication, alertness emotional behaviour Other: sleep/rest, eating, work, home management, recreational past times
4	McMaster Health Index (Chambers et al., 1982)	Physical: mobility, self-care, communication and global physical functioning Social: general well-being, work/social role performance, social support and participation and global social function Emotional: self-esteem, findings about personal relationships and the future, critical life events, and global emotional functioning
5	Nottingham Health Profile (Hunt, 1984)	Pain, physical mobility, sleep, emotional reactions, energy, social isolation
6	Short Form Health Survey (SF-36) (Ware and Sherbourne, 1992)	Vitality, Physical functioning, bodily pain, general health perceptions, physical role functioning, emotional role functioning, social role functioning and mental health.
7	European Quality of Life (EuroQoL) (R. Rabin, F. de Charro, 2001)	Mobility, self-care, usual activities, pain/discomfort and anxiety/depression

[Source: (Patrick and Deyo, 1989, Wenger et al., 1984, Cheung et al., 2009)]

Specific measures

Generic measures have been developed to meet the need for rapid classification of patients, and the sensitivity of these measures to small but clinically important change is probably limited (Patrick and Deyo, 1989). Condition-specific measures focus on a particular condition, disease, population or problem and are potentially more responsive to small, but clinically important, changes in health (Cunningham et al., 2000). The specific measures can further be classified as measures that are used to measure the General Oral Health Related Quality of Life and measures that are used to measure condition specific oral health related quality of life. Some of these measures are listed Table 1.3.

Table 1.3 List of Specific measures and their dimensions.

SNo	Measure	Dimensions measuring
Measures of General Oral Health Related Quality of Life		
1	Social Impacts on Dental Disease (Cushing et al., 1996)	Pain, difficulty with eating and communication problems
2	Geriatric Oral Health Assessment Index (GOHAI) (Atchison and Dolan, 1990)	Functional limitations and pain, psychological and behavioural impacts, discomfort, satisfaction, Self-consciousness,
3	Oral Health Impact Profile (Slade and Spencer, 1994, Slade, 1997)	Functional limitation, Physical pain, Psychological discomfort, Physical disability, Psychological disability, Social disability and Handicap
4	Oral Impacts on Dental Performance (OIDP) (Adulyanon and Sheiham, 1997).	Oral health and social and psychological factors
Measures of Condition Specific OHRQoL		
1	Instrument for Orthognathic treatment (Cunningham et al., 2000)	Social aspects of deformity, facial aesthetics, oral function and awareness of facial deformities.
2	Malocclusion Impact Questionnaire (MIQ) (Benson et al., 2016)	Uni-dimension

1.3.3 Measures for OHRQoL

It is widely accepted that many factors such as the patient's age, tooth loss, existing pathologies, sociodemographic, cultural, educational, psychological, dietary and financial factors contribute towards OHRQoL and involve different dimensions of oral health (John et al., 2004b, Kranjčić et al., 2014). The components involved in measuring OHRQoL (Figure 1.4) have been discussed by Inglehart and Bagramian (Inglehart and Bagramian, 2002). The OHRQoL of patients can be assessed by three distinct approaches. The three categories of OHRQoL measures stated by Slade (Slade, 2002) are social indicators, global self-ratings and multiple items questionnaires. Among these the multiple item questionnaires are particularly popular among researchers. Social indicators measure the oral condition in the community. These studies require a large population survey to measure social indicators such as number of days absent from work/ school due to dental problems, restricted social activities etc. Global self-rating is relatively fast and easy and gives a rough idea about the oral health quality of life of patients. This is a single item measure asking the individuals about their overall oral health in the past few days. The response to this question varies from very poor to very good. Though it is fast and simple to carry out, in multiple items questionnaires, more items produce replies that are more consistent and less prone to distortion from sociopsychological biases and this enables the random error of the measure to be cancelled out (Bowling, 2005).

To quantify OHRQoL, instruments with multiple items are required to represent the different dimensions. Moreover, the unknown constructs are complex and cannot be measured with a single item and therefore the use of multiple items provides more reliable quantification. The multi-item scales are more reliable and less prone to random measurement errors than single item measures

(Hahn, 2011). Hence, multiple item questionnaires which include items assessing different dimensions of oral health have been developed.

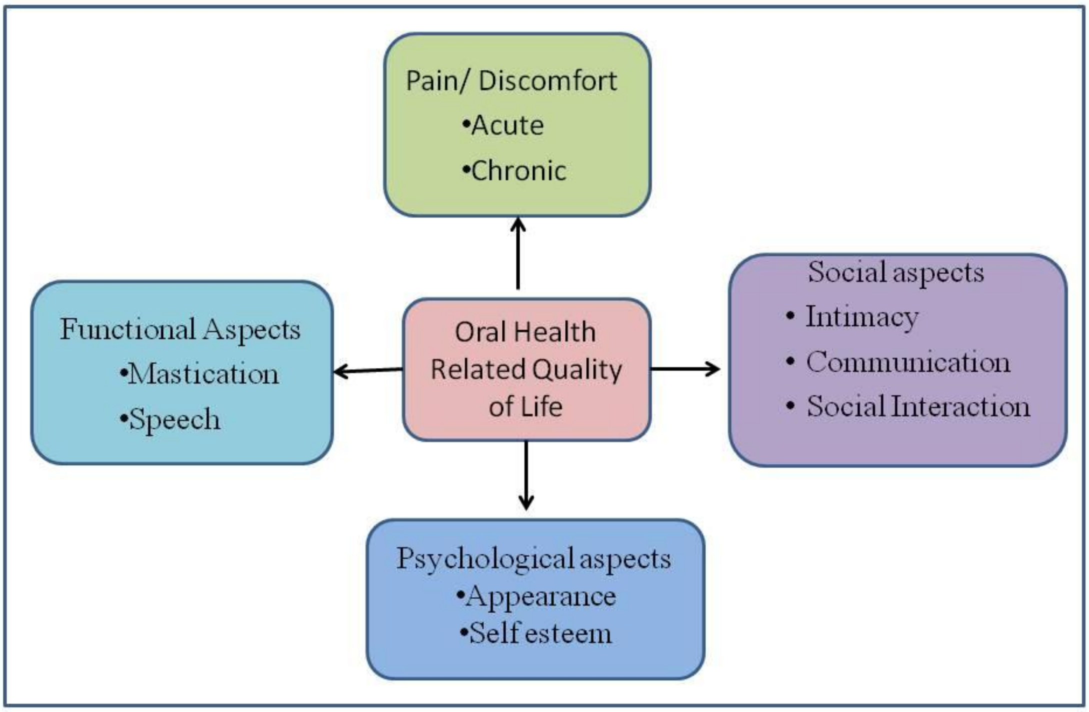

Figure 1.4 Components involved in measuring Oral Health Related Quality of Life

[Source: (Bennadi and Reddy, 2013)]

Though numerous instruments have been developed to measure OHRQoL (Skaret et al., 2004), each one differs in the number of items and dimensions. However, it is not clear which instrument is more appropriate for a study to measure OHRQoL. The number of questions in the instrument range between 3 (RAND Dental Health Index) to 56 (Oral Health Quality of Life Inventory) questions (Atchison and Dolan, 1990, Cornell et al., 1997, Leao and Sheiham, 1996, Locker and Slade, 1994, Slade and Spencer, 1994, Strauss and Hunt, 1993, Ware Jr and Sherbourne, 1992, Kressin et al., 1996, Kressin, 1996). At the First International conference on Measurement of Oral Health, a total of 10 instruments which were tested and evaluated for their psychometric properties were presented. The ten instruments along with the dimensions and other details are summarized in Table 1.4. (source: (Al Shamrany, 2006)).

Table 1.4 Oral Health Related Quality of Life Questionnaires

SNo	Measure	Authors	Dimensions measured	No. of questions
1	Socio Dental Scale	Cushing AM, Sheiham A, Maizels J. (1986)	Chewing, talking, smiling, laughing, pain, appearance	14
2	RAND Dental Index	Gooch, BE, Dolan, TA (1989)	Pain, Worry, Conversation	3
3	General Oral Health Assessment index	Kathryn A. Atchison Teresa A Dolan (1990)	Chewing, eating, social contacts, appearance, pain, worry, self- Consciousness	12
4	Dental Impact Profile	Strauss and Hunt, 1993	Appearance, eating, speech, confidence, happiness, social life, relationships	25
5	Oral Health Impact Profile	Slade and Spencer, 1994	Function, pain, physical disability, Psychological disability, social disability Handicap	49
6	Subjective oral health status indicators	Locker and Miller, 1994	Chewing, speaking, symptoms, eating, communication, social relation	42
7	Dental Impact on daily living	Leao and Sheiham, 1994	Comfort, appearance, pain, daily activities, eating	36
8	Oral Health Related Quality of Life	Kressin N, Spiro III A, Bosse R, et al. (1996)	Daily activities, social activities conversation	3
9	Oral impacts on daily performance	Adulyanon and Sheiham, 1997	Performance in eating, speaking, oral hygiene, sleeping, appearance, emotion	9
10	Oral Health Quality of life Inventory	Cornell, Saunders ,Paunovich,Frisch (1997)	Oral health, nutrition, self-related oral health, overall quality of life	56

Broder and co-workers (Broder et al., 2000), while studying the perceived impact of oral health conditions among minority adolescents, used two self-rated instruments: 1. RAND SF-36, an instrument with 36 items aiming to measure the health conditions under eight dimension namely physical functioning, bodily pain, role limitations due to physical health problems, role limitations due to personal or emotional problems, emotional well-being, social functioning, vitality, general health perceptions and perceived change in health; 2. Oral Health Impact Profile (OHIP) with 49 items (Slade and Spencer, 1994).

Mehta and co-workers (Mehta and Kaur, 2011) argued that multi item questionnaires capture more statistical variation than single item questions. Allen and Locker (Allen and Locker, 2002) used a short version of OHIP with 20 items and it has since been used by many other researchers (Allen et al., 2006, Awad et al., 2000, Ellis et al., 2008, John et al., 2004b). Awad and co-workers (Awad et al., 2014) have used OHIP-20 with a 6 point scale in a multicentre study in 8 countries to study the effect of mandibular 2 implant over-dentures on OHRQoL at baseline and 6 months after treatment and analysed the outcomes using seven dimensions (Slade and Spencer, 1994).

Locker and co-workers (Locker and Allen, 2007b) appraised five of the main instruments (GOHAI, OHIP, OIDP, COHQoL and OH-QoL) that have been developed over the past 20 years. They concluded that although there are different numbers of items in each instrument and they documented the frequency of functional and psychosocial impacts, but they failed to establish the meaning and significance of those impacts. But they all highlighted that the measures addressing quality of life should reflect the perspectives of patients rather than dentists.

1.3.4 Relationship between OHRQoL and demographic factors

Various instruments listed above have been used by many researchers to measure OHRQoL and established the relationship between OHRQoL and demographic factors such as age, gender, marital status etc. (John et al., 2004a, Slade, 1998, Steele et al., 2004, Wang et al., 2013). However, this relationship differed in studies and between populations. Studying the relationship between demographic factors and outcome measures is practically important in two ways. Firstly, for the prioritisation of treatment (treatment is targeted on the group whose OHRQoL is affected most) and second to measure the effect of treatment. In this endeavour, it is essential to know the current knowledge on the existing relationship so that it will allow to get an improveed estimate. .As this study is focussed on OHRQoL, the established relationship with demographic factors by some of the researchers has been discussed. In the Australian population, the effect of age and gender on OHIP-14 scores were small and statistically non-significant (Slade, 1997). Piovesan and co-workers found no association between socio economic factors and OHRQoL and concluded that there is no negative effect of oral health condition on children's quality of life in their study on school children (Piovesan et al., 2011). Separate studies carried out in Brazil and Sri Lanka identified the influence of socio economic factors on the poor perception of oral health

(Pattussi et al., 2007, Perera and Ekanayake, 2008). A strong association between the number of missing teeth and age, region of residence and income level has been reported in the Norwegian population (Åstrøm et al., 2006). They also reported that the prevalence of oral impact was comparatively lower in Norway compared to different OHRQoL measured in UK (McGrath and Bedi, 2002, Nuttall et al., 2001). Astrom and co-workers (Åstrøm et al., 2006) reported the poorest OHRQoL among lowest income group in the Norwegian population. A significant association between OHIP-14 score and general health was also obtained by Macedo and co-workers (Macedo and Queluz, 2011) in the Brazilian population. A list of selected published work on the relationship between OHRQoL and demographic and clinical factors is summarised in Table 1.5.

Table 1.5 Selected list of publications showing the relationship between OHIP-14 scores and demographic variables.

Author(s) and year	**Country**	**Instrument used**	**Factors assessed**
Slade, 1997	Australia	OHIP-49, OHIP-14	No. of missing anterior teeth*, No. of missing unreplaced anterior teeth*, No. of teeth with attrition*, No. of reported medical conditions*, Attended public dental clinic*, Perceived need for dental treatment*, Age (yrs)* and Born outside Australia.
Nuttall et al., 2001	UK	OHIP-14	Seven factors in OHIP-14
Sheiham et al., 2001	UK	OIDP	Age, gender, social class* and region*.
McGrath and Bedi, 2002	UK	OHRQoL-UK	Age*, gender, social class* and number of teeth possessed*
John et al., 2004	Germany	OHIP_German Version	Denture status*, age, gender, education and residential area*.
Steele et al., 2004	UK	OHIP-14	Age* and number of remaining teeth*, sex, denture wearing* (partial or complete) and area*.
Astrom et al., 2006	Norway	OIDP	Age*, region*, dental attendance* and tooth loss* and income.
Pattusi et al., 2007	Brazil	Custom made and clinical examination	Social class*, sex*, ethnicity, Social support*, family structure*, Behavioural problems*, self-rated health*, dental attendance*, untreated dental carries*, missing teeth*, periodontal disease, dental pain*, mouth appearance* and chewing function*.
Walter et al., 2007	Canada	OHIP-49	Education, last dental visit, age, gender, no. of natural teeth, natural anterior missing, natural posterior missing, fixed partial dentures, removable dentures, implant borne fixed crowns, implant supported over dentures, treatment needs, gender, anterior missing*, treatment need in endodontics*
Perera and Ekanayake, 2008	Sri Lanka	Custom made	Gender, ethnicity, income*, use of dental services*, tooth brushing*, decayed teeth, missing teeth, gingivitis*, awareness*, tooth ache*, oral symptoms* and perceived need*.
Macedo and Queluz, 2011	Brazil	OHIP-14	Level of education, occupation, mean frequency of tooth brushing (number of times *per* day), general health*, dry mouth, smoking, gingival

Author(s) and year	Country	Instrument used	Factors assessed
			bleeding* and use of medications for toothache relief.
Zhou et al., 2012	China	OHIP-14	Gender, Age, Education*, Income and denture status*

* Denotes the factors that were significantly associated with OHRQoL

1.4 Oral Health Impact Profile -14 (OHIP-14)

Initially, the Oral Health Impact Profile (OHIP) was developed by Slade and Spencer by incorporating 49 items describing the consequences of oral disorders, derived from 535 statements obtained in interviews with 64 dental patients. This was then tested in the Australian population (Slade and Spencer, 1994). Later, the shortened version with 14 items was developed (Slade, 1997). The initial testing of the instrument by Slade was only by carrying out the secondary analysis of data which were collected in 1991-92 in the Australian population. The necessity to shorten the long form of OHIP has been discussed by Locker and Allen (Locker and Allen, 2002), who identified four major issues in using the long form which included - time consumption and question of feasibility in clinical settings, cost of administration and data management, burden on the respondents and the problem of high non response rates or increased missing data.

OHIP-14 was shown to have similar psychometric properties to the original version OHIP-49 (Slade, 1997). As this questionnaire is convenient and easy for patients to complete, it has become increasingly popular among researchers for both population based and clinical research projects (Ekanayake and Perera, 2003, Oliveira and Nadanovsky, 2005, Saub et al., 2005). Any new instrument has to be tested for its reliability and validity in the set up that is going to be employed. The reliability of an instrument evaluates the internal consistency, inter-rater reliability and stability of measures whereas the validity is the extent to which the interpretation of the results is acceptable.

OHIP-14 has been translated into various languages (Chinese, Japanese, German, Arabic, Croatian, Spanish, Malaysian, Dutch, Brazilian, Portuguese, Hungarian and Korean) (Bae et al., 2007, Szentpétery et al., 2006) and has been tested in various populations (Montero-Martín et al., 2009, Ziętek and Malicka, 2015, Motallebnejad et al., 2011) for its reliability and validity. Montero-Martin and co-workers translated it in to Spanish and tested in the Spanish population. Their study showed a Cronbach's alpha of 0.89 and they concluded that OHIP-14 is a reliable and sensitive measure to assess the OHRQoL in the Spanish population. Motallebnejad and co-workers (Motallebnejad et al., 2011) developed a Persian version of OHIP-14 which showed a Cronbach's alpha score of 0.95. Baker and co-workers (Baker et al., 2006) used OHIP-14 in Xerostomia patients in the UK and concluded that it has good internal reliability, good criterion and construct validity and that OHIP-14 performs better than OIDP. The suitability of OHIP-14 for assessing OHRQoL and its responsiveness was tested by Baba and co-workers (Baba et al., 2008). They used the Japanese version of OHIP-14 and OHIP-5 and reported that OHIP-14 had a Cronbach’s alpha of 0.94 and concluded that responsiveness of OHIP-14 was sufficient. Ikebe and co-workers (Ikebe et al., 2012) showed that OHIP-14 and Geriatric Oral Health Assessment Index were highly correlated and OHIP-14 had higher internal consistency (Cronbach's alpha=0.95). Robinson and co-workers (Robinson et al., 2003) compared OHIP-14 with OIDP and suggested better validity and reliability measures for OHIP-14, and concluded that OHIP-14 is superior to OIDP based on the data collected in the UK population. They reported a Cronbach's alpha value of 0.91 for OHIP-14 which indicates high reliability of this instrument and suggested that OHIP-14 is more suitable for questionnaire based research and for comparing groups while assessing the OHRQoL.

It is common in studies using multiple item questionnaires to check the presence of unknown constructs which represent various dimensions. Many researchers have attempted to study the number of unknown constructs in OHIP-14 items in different populations. A study carried out by Brennan and co-workers (Brennan and Spencer, 2004) in the Australian population, measured and compared the dimensions of OHRQoL measured by EQ-5D+ (Euro QoL) and OHIP-14. In this study they identified that OHIP items tended to load heavily on the first two factors indicating the possibility of only two factors in OHIP-14. The overlap of items both in EQ-5D+ and OHIP-14 was particularly for pain. The authors also raised the technical aspects of factor analysis as the items were not separated out. This is mainly because the correlation or covariance analyses require data of interval nature. Even in Slade's (Slade, 1997) original paper in deriving the short form of OHIP, they found one principal component accounted for the large proportion of variance which lead to the least satisfactory results from the factor analysis, which they attributed to the moderate to high correlation of all items to the overall score.

While using OHIP-14 to measure OHRQoL, the data were analysed either by considering the total number of impacts (number of items scoring above a particular impact threshold) or the mean score/ the total score based on all 14 items and the relationship with the demographic variables were studied using this. A weak association between age and mean OHIP-14 scores among people who had none of the five clinical symptoms listed (5 or more missing teeth, denture, untreated decay, moderate to severe periodontitis and tooth ache) has been reported (Slade and Sanders, 2011). However, a threefold association was found between age and mean OHIP-14 scores among adults who had 2 or more clinical conditions. It was also observed that 84.9% of dentate adults had OHIP-14 scores greater than or equal to 1. In this study, the mean OHIP-14 scores significantly differed by gender, country of birth, income and reason for dental visits. Interestingly, the mean OHIP-14 score is significantly higher among low income groups. Queiroz Herkrasth and co-workers in their meta-analysis, have found that there is statistical heterogeneity among studies with respect to the overall scores. Poor investigation of socio-economic and demographic factors have been identified in Oral Health research by Queiroz Herkrasth and co-workers (Queiroz Herkrath et al., 2015).

Since its introduction, OHIP-14 has been used extensively in a wide range of settings and different populations. It is notable that there are a wide range of total scores found in some of these different studies, and reach differing conclusions on OHRQoL. For example, OHIP-14 has been used by Macedo (Macedo and Queluz, 2011) to identify the self-perceived OHRQoL among workers in the furniture industry. This study observed the average OHIP score of 0.51 based on 111 samples. No significant association was found between demographic variables (age, gender and level of schooling). Guzeldemir and co-workers (Guzeldemir et al., 2009) used OHIP-14 and GOHAI while assessing the OHRQoL of patients undergoing haemodialysis and found that there is a significant correlation between OHIP-14 and GOHAI. They identified that OHIP-14 had more impact on OHRQoL than GOHAI. The mean OHRQoL of haemodialysis patients was 19.40 with a standard deviation of 7.74. A positive relationship between age and OHIP-14 scores was established by Jain and co-workers (Jain et al., 2012) in an Indian Population. They used OHIP-14 to study the impact of Age and tooth loss on OHRQoL in two states namely Gujarat and Rajasthan in India. The mean OHIP-14 scores (in 5 point scale) were 7.05 and 10.40 for Gujarat and Rajasthan respectively indicating regional differences. OHP-14 was also used to study the order effect and influence of reference period on OHIP summary scores (Sutinen et al., 2007). This study reported the absence of an order effect while using OHIP-14 for studying the two reference periods. Though a one year reference period has been widely accepted, other reference periods (two weeks, one month, 3 months) have also been used. While a one year reference period may be suitable for national surveys, shorter reference periods are more suitable for studies that deal with the short term effect of oral conditions. Sutinen and co-workers (Sutinen et al., 2007) have reported that the mean OHRQoL score for patients who were waiting for orthognathic treatment was 16.8 while for the control group it was 4.3. These measures were obtained by using OHIP-14 with a one month reference period. The positive relationship between OHIP-14 total score and the presence of oral disease and inverse relationship between OHIP-14 total score and age has been reported (Robinson et al., 2003).

1.5 Methodological Challenges of questionnaire based data

Most patient-reported outcomes are measured using questionnaires. The questionnaires are usually developed with items whose answers are of Likert scale type ranging from 0 to 5. In such types of measurement, problems can arise not only in the method adopted to quantify what we intend to measure but also in the method of analysing it. Some of the common issues in such studies include missing data, floor and ceiling effects, responsiveness of the measure and the identification/ confirmation of a number of unknown constructs (dimensions) in the items that are used during the quantification process. Identification/ confirmation of numbers of dimensions can be considered as one of the methodological issues as most of the instruments are based on some theoretical framework and hence the items are supposed to represent the various dimensions of the model.

1.5.1 Missing data

In questionnaire based data, it is common to have data missing in one or more items. The conventional method of analysing datasets with missing data is to ignore the missing cases and then the analyses are based on completed cases only (Briggs et al., 2003, de Goeij et al., 2013). But the reliability of the results obtained based on the complete case analyses depends on the amount of missing data in the study and the pattern (mechanism) of missing data in the dataset, along with the size of the sample. Methods of handling missing data vary but generally involve imputation by techniques such as item mean, subject mean, regression, interpolation, Expectation Maximisation (EM) algorithm and multiple imputation. The use of a multiple imputation method has been recommended by Royston and co-workers (Royston, 2004) for handling missing data. However, the choice of method depends on the mechanism involved in missing cases.

1.5.2 Floor and Ceiling effects

The majority of the outcomes measuring the social and psychological impact of treatments have been defined by Likert scales, where items are scored within a finite range of values, often with five or six possibilities - 0 to 5, where 0 indicates no impact and 5 indicates the worst impact. It is common in questionnaire based data collection, particularly in surveys based on Likert scales, that many of the respondents tend to give the lowest score or the highest score as there is no other choice to rate higher or lower. This type of scale, and any derived composite (the total score of the items forming a given dimension, for example), yields measures that are ‘censored’ on the right, on the left, or on both sides. In statistics, censoring occurs when the value of an outcome is only partially known. As an example, if a patient experiences a severe day-to-day problem due to illness, they are forced to give a score of 5 (the maximum) while measuring the Quality of Life. If the range of the score had been from 0-100 instead of 0-5, the individuals who were forced to give 5 for a particular question would have given different scores in the higher range. By adopting a shorter range (0-5), the items are ‘right-censored’, meaning that the true score of these individuals may not have been measured accurately but have been curtailed at the highest value recorded. This curtailment may be drastic and any bias introduced in this manner is termed the *ceiling effect*. This potential bias arises because the variability between those individuals with severe dental problems cannot be measured. Likewise, when the items are ‘left-censored’, we say that a ‘floor’ has been fixed by the score and the bias introduced in this manner is termed the *floor effect*. As the participants who have given lowest / highest scores could possibly be further divided into different categories, had they been given more choices to rate their severity of illness, this will affect the relationship between the outcome and other predictor variables. A ‘ceiling’ occurs when the participant is forced to give a maximum, so called the ‘ceiling’, which is much less than the asymptotic value (a value that is close to the actual value). The presence of Floor and Ceiling effects in Likert scales has long been established (Oord and Ark, 1997) and they proposed a method to calculate these effects in the data. The study by Astrom and co-workers (Åstrøm et al., 2005) using the Norwegian version of Oral Impact on Daily Performance Inventory (OIDP) showed a total of 81.7% had zero OIDP scores and 30.0% had the maximum score.

In the presence of censoring, estimators of the effects based on the conventional Ordinary Least Squares (OLS) models are inconsistent (in the sense that the coefficients will not necessarily approach the "true" population parameters as the sample size increases) and yield biased estimations of both the gradient (under-estimating it) and the intercept (over-estimating it) of the model. Many researchers have studied the impact of the ceiling effect. Some studies have identified artificial non linearity and under estimation of the regression parameters (Macedo and Queluz, 2011, Slade and Sanders, 2011). At the same time, researchers have developed new analytical techniques that can handle the ceiling effect. The use of growth curve models have been suggested for handling the ceiling effect in longitudinal data (Oord and Ark, 1997, Wang et al., 2008). In a cross sectional dataset, Tobin (Tobin, 1958) developed the Tobit regression model, based on econometrics data, which has been widely used in cross sectional studies where the measure has a ceiling effect. These approaches apply to both ceiling and floor effects in a similar way. Many researchers have attempted to handle the floor and ceiling effects in various measurements. Brown and Muthen (Brown, 1989, Muthén, 1989) have separately documented the benefit of using Tobit regression to address the shortcomings of Pearson correlation when analysing censored data. Therefore, any model fitted without adjusting for floor and ceiling effects, when a considerable proportion of cases suffer from these would be underestimating (in terms of magnitude) the effects. Based on simulated data, McBee (McBee, 2010) also established similar results. Wang and co-workers (Wang et al., 2008), using empirical and simulated data, showed that ceiling effects in longitudinal data lead to biased parameter estimation.

1.5.3 Factorial structure

As most of the studies are based on self-rated questionnaires (Patient Reported Outcomes, PROMs) involving various items related to different activities of the patients in their day to day life (multidimensional), it is essential to assess the construct and discriminant validity of the instrument by confirming the number of unknown constructs in the instrument used. In this process, factor analysis is a useful tool to investigate the relationships between variables for complex concepts that are not easily measured. The need to check the factorial structure of the instrument has been stressed particularly if it is developed based on some theoretical models in order to assess its construct validity (Lo Coco et al., 2008). The construct validity is mainly evaluated through factor analysis and/or Structural Equation Modelling (SEM). Moreover, this helps to identify the meaningful subscales within the items in the instrument.

1.5.4 Responsiveness

Though many instruments have been used to assess OHRQoL (Slade, 2002, Al Shamrany, 2006) and many studies have attempted to quantify OHRQoL, very few have attempted to measure the responsiveness of the instrument used (Sutinen et al., 2007). While measuring the responsiveness, there is no consistency between studies regarding the reference period and different studies have used different reference periods. Sutinen (Sutinen et al., 2007) used one and twelve months reference periods, whereas a study carried out in India by Kavin and co-workers (Kavin et al., 2012) used 2 months and 6 months reference periods. As complex dental problems may need a longer time period to heal, a longer reference period is required to measure the actual effect of treatment when compared to that for simple dental problems.

As change in health status is an important measure to assess the treatment effect or the effectiveness of any intervention programme, many researchers have worked on this issue (Guyatt et al., 1987, Kazis et al., 1989, Locker, 1992, Locker et al., 2004, McGrath and Bedi, 2002, Ziebland, 1994, Deyo et al., 1991). Two types of changes exist while measuring health status, namely, qualitative and quantitative (Menard, 1991). Qualitative change is whether there is a change in status of an individual over the time period or not, whereas the quantitative change is analytically more difficult (Locker, 1998). Streiner (Streiner and Norman, 1989) has explained three analytic tasks in measuring change. The first one is to identify the change and the difference between individuals, groups etc. The second one is to identify the significant predictors of change and lastly to explain the change as such. It is difficult to answer the differences in change between individuals and to decide whether the change is due to treatment or due to some other factors. Ziebland (Ziebland, 1994) has clearly explained the four different ways in which the change scores can be measured. They are:

1. Comparison of distributions or comparing scores at baseline and follow-ups. In this, the score at baseline and follow up are compared for any change in status.

2. Use of change scores, which provides a quantitative approach for the analysis.

However, this leads to two different problems:

1. Definition of clinically meaningful change and
2. Psychometric properties of change scores.

The problems of clinically meaningful change can be addressed either by calculating the effect size and standardised response means or by correlating health status measures to change scores derived from clinical measures (Deyo et al., 1991, Jenkinson et al., 1997). The problem of psychometric properties of change scores can be addressed by measuring the reliability coefficient. Streiner and Norman (Streiner and Norman, 1989) suggested that change score may be used if the reliability coefficient is greater than or equal to 0.5.

3. Global transition judgments - This method allows the patients to self-rate their oral health status by a single overall score. In this method patients will be asked whether there is a change in their oral health status during the reference period. As an example, patients may be asked "Over the past four months your oral health has 1. Improved 2. No change or 3. Worsened." However, the use of this method to measure the change has been criticised (Ware Jr et al., 1981). This method may fail to differentiate the changes in various dimensions – improvement in some dimensions and deterioration in others (Ziebland, 1994). Some researchers have

established that there is a close association between overall score with multi-item/ multi-dimensional scores (Rowan, 1994).

4. The fourth measure can be derived from a series of global transition judgments from different dimensions of oral health. This can be measured by using Likert scale type of individual items. For oral health, the transition scale might be obtained from asking various questions related to chewing, appearance, pain, worry and social activities during the reference period. The transition scores are then calculated by totalling the individual question score which may be answered using Likert scale type of responses. Researchers have reported that such transition scores are the better measure of change (Ziebland, 1994).

Each of the above four ways of measuring change has its own advantages and disadvantages. Therefore, none has been universally accepted and the choice of the best method remains controversial. Among these, the use of change scores (the difference in scores between pre and post intervention period) is commonly used in dental research (Allen et al., 2009).

Need to address methodological issues in Oral health research

As OHRQoL depends on cultural and social norms, the cross cultural relevance of the measure has to be explored. In this regard, the importance of national norms in measuring OHRQoL for national surveys has been realized (Allen, 2003). The quantification of oral health mainly depends on the type of questionnaire used to measure and the items involved in it. Allen has stressed the need for methodological work for measuring sensitivity to change properties (Allen, 2003). The need to address methodological issues in oral health research has been stressed by Inglehart and Bagramian (Inglehart and Bagramian, 2002) and they have highlighted the development of outcome measures for longitudinal studies, appropriateness of measures as influenced by the passage of time, sensitivity, specificity, reliability, and validity as some of the issues.

1.6 Methodological issues in OHIP-14 data

Though many studies have attempted to quantify OHRQoL, few have attempted to explore the methodological issues that arise in measuring it. The issues while measuring OHRQoL using OHIP-14 are discussed in detail below.

1.6.1 Missing Data

As discussed earlier, the problem of missing data is common to many studies, particularly when based on self-rated questionnaires. Measuring OHRQoL using OHIP-14 is no exception. To get an efficient estimate with good power, the missing data pattern has to be explored and suitable methods to analyse the data need to be devised. However, most studies carried out to assess OHRQoL fail to report the missing data in their data set. Some studies (Corridore et al., 2013, Hongxing et al., 2014, Motallebnejad et al., 2011, Rusanen et al., 2009) have reported the missing data, but still failed to address the handling of missing data while carrying out the analysis. . Locker and co-workers (Locker, 1998) have listed attrition as one of the methodological problems in measuring the change scores. Slade and Sanders (Slade and Sanders, 2011) have identified missing data in the items while measuring the OHRQoL. However, they have not explored the missing data mechanism but have handled missing data by substituting with item mean values. Locker (Locker, 2004) studied the responsiveness of OHIP-14 by measuring pre and post treatment OHRQoL scores. In their study, they encountered an attrition rate of around 44% but neither the missing data mechanism was explored nor was a suitable method used to handle missing data. While validating the Hungarian version of OHIP-14, Szentpe´tery and co-workers (Szentpétery et al., 2006) excluded the subjects from the analysis if the summary scores contained any missing data. For subjects having missing data, they imputed using a regression approach. But the missing data mechanism was not explored. Lahti and co-workers (Lahti et al., 2008) have dropped cases from the analysis if they contained more than two items missing which accounted for around 6% of the total sample. Guarnizo-Herreño and co-workers (Guarnizo-Herreño et al., 2014) analysed Adult Dental Health Survey 2009 (ADHS 2009) data for the socio economic position and oral health in UK. In this study, though they reported the missing data, the analyses were carried out based on completed cases and the missing data mechanism was not explored. For the purpose of assessing the effect of missing data based on regression models, the missing data were imputed using multiple imputation and simple regression methods. Saub and co-workers (Saub et al., 2005) while developing the short version of Malaysian OHRQoL measure using the subsample of participants from the Malaysian National Oral

Health Survey of Adults (NOHSA, 2000), handled missing data by two approaches namely total exclusion and imputation. Total exclusion was adopted for cases with more than 20% of the items missing and item mean imputation was carried out for cases with missing items less than or equal to 20%. They achieved an overall response rate of 50.9% and an incomplete data of 1.5% among the completed cases. Out of 206 respondents who completed the initial questionnaire, only 73 completed the second copy after a two weeks interval leading to a response rate of only 35.4%. The response rate of 32.5% was achieved for the mail questionnaire after two weeks interval from the first data collection. Awad and co-workers (Awad et al., 2014) reported around 50% of missing data in their study. The missing data were imputed by multiple imputation, without exploring the missing data pattern. None of the socio demographic variables showed significant differences between complete and incomplete groups. Also they emphasised the need for handling missing data in studies where the missing data is encountered and suggested the need to use the complete data for analysis. Baba and co-workers (Baba et al., 2008) while testing the suitability of OHIP-14 in a Japanese population, handled missing data by imputing them with regression methods for few missing items and the responses were ignored from the analysis if more items were missing (more than 5). In a recent study carried out in Australia to find out the predictors of dental avoidance among different levels of dental anxiety using OHIP-14 by Armfield and co-workers (Armfield and Ketting, 2015), an adjusted response rate of 41.1% was achieved. They reported that 48% of the cases contained incomplete data and the item missingness ranged between <1% to 21.8%. The missing values were imputed using a multiple imputation technique. The

Japanese version of OHIP-49 was used in a follow up study carried out by Fueki and co-workers (Fueki et al., 2015) to study the effect of prosthetic treatment in the Japanese population. The follow ups carried out at 3, 6 12 months interval from the baseline showed that 26.0% were lost to follow up in the post treatment evaluations. The missing data were ignored in this study and complete case analyses were carried out. Reissman and co-workers (Reissmann et al., 2013) reported only 0.47% of all items in OHIP were missing in their study on the impact of donor site for bone graft harvesting for dental implants on Health-Related and OHRQoL. The missing data in this study were imputed with regression methods. Gisler and co-workers (Gisler et al., 2012) reported 5.5% of respondents had missing items and these cases were excluded from the analysis while studying the relationship between dental anxiety and OHRQoL using the German version of OHIP-14 in the Swiss population. Although Allen and co-workers (Allen et al., 2009) observed that around 14% out of a total sample of 51, failed to follow up, the missing data were not taken into account and complete case analyses were carried out.

Although the above researchers have discussed the missing data in their studies or imputed with mean values while measuring OHRQoL using OHIP-14, none have explored the missing data mechanism in their data set. As the choice of analysis depends on the missing data mechanism, it is important to explore the missing data mechanism in OHIP-14 while measuring OHRQoL. As the analysis depends on the missing data mechanism and a considerable amount of missing data has been reported when using OHIP-14, the missing data mechanism has to explored while measuring OHRQoL using OHIP-14 and an appropriate analysis strategy has to be devised.

1.6.2 Floor and Ceiling effects in OHIP-14 data

As the responses to OHIP-14 are from 0 to 4, where 0 represents “never” and 4 represents “very often”, there is a high chance of recording extreme values (either 0 or 4). In OHIP-14, if the responses to an individual item vary from 0 to 4, then the item ceiling takes place at 4 and the ceiling for the total score (the total scores of all 14 item) will be at 56. All those individuals with severe dental problems will score 4 for an item, the maximum allowed. Many studies have analysed the relationship between OHIP-14 scores and demographic variables although none seem to have taken floor and ceiling effects into account. Similarly, the presence of floor effect of the OHIP may limit the ability of the questionnaire to measure intra individual changes in OHRQoL in older adults (Locker et al., 2001).

The presence of a high rate of floor effect in OHIP-14 has been reported by Locker and co-workers (Locker et al., 2001). They reported around 30% of the respondents scored 0 in the additive method and 45.8% in the count method. In the additive method, the scores of all the 14 items were summed up to get a single total score (ranging between 0 to 56: 14 items each with a minimum of 0 and maximum of 4) whereas in the count method the count scores are created by counting the items with responses ‘‘sometimes’’, ‘‘fairly often’’, ‘‘very often’’ or ‘‘never’’ (range between 0 to 14). Ikebe and co-workers (Ikebe et al., 2012) showed that 12.1% of cases had 0's in OHIP-14 score whereas only 4.6% had zero score in GOHAI score. Locker and co-workers (Locker et al., 2001) found that OHIP-14 had large number of 0's (30.3%) indicating greater floor effect when compared to GOHAI (8.4%). Similarly, Hassel and co-workers (Hassel et al., 2010), while studying the oral health of Germans showed 34% had 0's in OHIP-14 score while only 7.3% in GOHAI score. Ikebe and co-workers (Ikebe et al., 2012) concluded that the presence of a high floor effect may limit its use in longitudinal studies. However, none of the studies have taken the floor effect into account while analysing their data or attempt to address the floor effect in detail.

A review of the literature has shown that though many studies have stated the presence of Floor and Ceiling effects while measuring OHRQoL using instruments particularly OHIP-14, based on Likert scale, none has incorporated these effects in the analysis. This leads to under estimation of various parameters and affects the precision of the analysis results.

1.6.3 Factorial structure of OHIP items

Another important issue in measuring OHRQoL is the identification of the unknown constructs in the instruments used. The seven dimensions identified by Slade (Slade, 1997) using both the short and long version of OHIP were Functional limitation, Physical pain, Psychological discomfort, Physical disability, Psychological disability, Social disability and Handicap. Although the existence of seven dimensions has been widely accepted, different researchers have proposed different numbers of factors in OHIP-14. The study by Short and co-workers (Short and Horn, 1984) suggested four variables per factor and reported that if the factors are measured using a lesser number of variables (less than 3), then the factors cannot be well located in the objective rotation and the factors will not be reliably distinguished in the analysis (Horn, 1967, Horn and Knapp, 1974, Horn and McArdle, 1980, Humphreys et al., 1969). Accordingly, as OHIP-14 contained only 14 items, considering more than 3 factors (each factor with 4 items) will not be able to distinguish the factors reliably. An improved factorial structure may be obtained by using the original OHIP-49 which has 49 items to measure the OHRQoL. But, the original study by Slade and Spencer (Slade and Spencer, 1994), found that one principal component accounted for a large amount of the variance whereas another three factors had eigen values greater than 1 (Eigen values measure the amount of variation in the total sample accounted for by each factor and is used to decide the number of factors in the factor analysis (Fruchter, 1954, Norris and Lecavalier, 2010)). They also suggested that the higher inter correlation between the items may be the reason for a single principal component contributing to a large percentage of variance. This suggested that a single underlying construct could represent the oral health of an individual. Therefore, the confirmation of the number of factors in OHIP-14 is necessary to assess the construct and discriminant validity of the instrument and to facilitate further factor wise analyses of data.

Baker and co-workers (Baker et al., 2008a) while testing the construct validity of the Oral Health Impact Profile as a measure of Locker's conceptual model using

Structural Equation Modelling (SEM), reported that the data did not support the construct validity of the OHIP-49. They also highlighted that the error terms are highly correlated and many items happened to measure more than one construct. They arrived at a six factor model comprising 22 items of OHIP-49. In the same work, while exploring the relationship between the factors conceptualized by Locker's model, they found only a partial relationship between the factors. They suggested that OHIP sub-scales should not be reported and different dimensions of OHIP should not be distinguished while analysing treatment effect.

The qualitative analysis of oral health by Gregory and co-workers (Gregory et al., 2005) identified seven dimensions of oral health relating to the positions that people would adopt. The seven dimensions arrived by Gregory and co-workers are,

1. Positioning of the norm (Health)
2. Positioning of attribution (Internal)
3. Positioning of Dentistry (Trust)
4. Positioning of accessibility (Choice)
5. Positioning of commodity (Embracing)
6. Positioning of Authenticity (Natural)
7. Positioning of Character (Admiring)

in constructing the oral health relevance. In this paper, the authors have stated that the meaning of quality of life changes over time and this was also supported by Barbosa and co-workers (Barbosa and Gavião, 2008). Donnelly and co-workers (Donnelly and MacEntee, 2012) quoting Gregory's work, defined oral health as a dynamic phenomenon influenced by many factors that change with time and circumstances. Passalacqua and co-workers (Passalacqua et al., 2012) in their efforts to assess oral health promotion programmes in UK, highlighted the importance of oral health knowledge in achieving good oral health and evaluated the four dimensions (trust, natural, admiring and choice) of Gregory and co-workers (Gregory et al., 2005). Another qualitative study by Gregory and co-workers (Gregory et al., 2005) states that people construct their own 'margins of the relevance' of oral health. If the relevance of oral health differs between individuals then the importance and role of the seven dimensions defined by Gregory and co-workers (Gregory et al., 2005) is of practical importance for planning and evaluating oral health programmes.

The discriminant validity of an instrument is assessed by discriminating different subscales from one another within the multidimensional measurements of an instrument (Ong and van Dulmen, 2007). The dimensional validity is important because it indicates how an instrument should be used in practice (Santos et al., 2013). Therefore, dimensional and construct validity of OHIP-14 is checked by taking four different models that are in use to measure OHRQoL using the study data set.

1.6.4 Responsiveness to change

The measurement of change in OHRQoL is pertinent to any study that attempts to assess the change in OHRQoL of patients over a period of time. The measurement of change in health status is more complex as well as controversial (Locker, 1998) and hence a challenge to researchers. The responsiveness to change is an important issue to be explored as we need to understand the impact of therapeutic interventions on the quality of life (Allen, 2003). Tu and co-workers (Tu et al., 2005) have clearly explained the misuse of correlation between pre-treatment and post treatment values while finding the effect of intervention in any disease. As the post treatment value is contained within the pre-treatment value, this leads to mathematical coupling (Archie Jr, 1981, Anderson and Dedrick, 1990). Allen (Allen, 2003) has stressed the need for methodological work for measuring sensitivity to change properties.

Locker and co-workers (Locker, 2004) while examining the sensitivity of OHIP-14 to change have highlighted that where a measure has proven to be valid and reliable in cross sectional studies it cannot be assumed that it will be suitable to assess the outcome of clinical intervention. A significant change in oral health is measured as the minimally important difference required. The commonly used methods for measuring the Minimal Important Difference (MID) are effect size and Standard Error of Measurement (SEoM) (Revicki et al., 2008, Tsakos et al., 2012, Wyrwich et al., 2013). Locker (Locker, 2004) used change score and Guyatt's responsiveness statistic to measure the change in OHRQoL in the Canadian population. Locker considered pre-treatment and one month post treatment period for measuring the change. The mean OHIP-14 score declined from 15.8 to 11.5 and the decline was statistically significant. Among the 116 samples, 60.2% reported improvement during the post treatment period. They also found that OHIP-14 appeared to be responsive to change but the magnitude was modest. The effect size based on all subject was 0.32. They also tested the change using paired t-test, testing the difference between pre and post treatment scores. This showed that the change in OHIP score was prominent among the group who reported improvement but change was not significant in the stable group. Locker also used ROC analysis with various cut of points for OHIP-14 scores while measuring the change in OHRQoL.

Durham and co-workers (Durham et al., 2011) have developed a short form of OHIP-49 for Tempero-Mandibular Disorders (TMO) with 20 items from OHIP and two new items totalling 22 items. The psychometric properties and the effectiveness of this instrument to measure the responsiveness to change has been tested by Yule and co-workers (Yule et al., 2015). The time period between the two data collection points was two weeks in this study. The response rate from the first to second data points was 54.7%. They used effect size as a measure of responsiveness to change, which was 0.4 for both OHIP-49 and the shortened version. The minimally important clinical difference was assessed using the mean change scores.

The two main factors that influence the change in OHRQoL are the instrument used and the reference time period used in the study. Usually in a longer reference period, the change may be obvious when compared to shorter periods. Lodi and co-workers have suggested the need for lengthy follow up periods (Lodi et al., 2012) to evaluate the long term benefit. A longitudinal Randomised Controlled Trial aimed to measure the plaque control intervention in the UK population by Stone and co-workers (Stone et al., 2015) used OHIP-49 for measuring the effect at two follow ups at two and 20 weeks interval and reported a loss of 3 subjects during the follow up. Similarly, the presence of redundant items, which may dilute the responsiveness to change (Yule et al., 2015) in OHIP-49 while measuring the change, has been widely discussed by many researchers (Durham et al., 2011, John et al., 2002, Larsson et al., 2004, Segù et al., 2005). Kimura and co-workers (Kimura et al., 2012b) highlighted the need for future studies to measure the response shift in different populations and using different instruments like OHIP.

The sensitiveness and responsiveness of OHIP have been studied by many researchers. The study by McGrath and co-workers (McGrath et al., 2005) using OHIP-49 reported results similar to Locker and co-workers (Locker et al., 2004). The overall effect size observed in this study was 0.31 indicating the magnitude of change was moderate; but, the change was apparent in overall OHIP and hence the researchers have concluded that OHIP is sensitive and responsive to change in OHRQoL. However, these studies have used mean change in OHIP scores and effect size as a measure of change/ responsiveness. The changes calculated in these studies were from the data collected using OHIP-49 and the post treatment measures were taken after 8 weeks. Therefore, the suitability of the shorter version of original OHIP (OHIP-

14) in measuring responsiveness to change still needs to be explored. Similarly, the efficiency of OHIP-14 in measuring the change scores in OHRQoL has not been well established.

1.6.5 Chapter Summary

The review of the literature on the use of OHIP-14 in measuring OHRQoL has highlighted that though OHIP-14 is more common among researchers in measuring OHRQoL, the data collected using this are subject to some of the methodological difficulties identified. Though missing data is common in many studies and some have highlighted this issue, none have attempted to explore the missing data mechanism and none have used a suitable method to address it based on the mechanism. Similarly, though some studies have mentioned the presence of floor and ceiling effects (Hassan et al., 2017, Ikebe et al., 2012, Wright et al., 2009), none have attempted to study the impact of these effects in predicting the OHRQoL. The effectiveness of OHIP-14 in measuring the responsiveness to change in the presence of missing data has also not been explored adequately. The number of dimensions in OHIP-14 items varied from 1 to 7 for different studies. These issues are to be explored in detail and suitable analyses strategy to be devised to address these issues so that the OHRQoL can be quantified more precisely in oral health research. Therefore, the present study has attempted to address these issues with the following objectives.

1.7 Aims and Objectives

Aims

The two main aims of this study are:

- To evaluate the methodological issues that arise while measuring Oral Health Related Quality of Life using OHIP-14.

- To evaluate the responsiveness (sensitivity to change), internal consistency and the inter item correlation of OHIP-14.

Objectives

- Quantify the OHRQoL of a random sample of all patients attending King's College London Dental Institute, King’s College Hospital Campus, UK for their dental treatment.
- To explore the missing data mechanism in collecting oral health data using OHIP-14 and to compare various missing data handling methods.
- To evaluate the bias introduced by the presence of missing data in longitudinal studies of OHRQoL and to find a suitable method to treat missing data in OHIP-14.
- To identity the floor and ceiling effects and to incorporate the effect of these findings on the relationship between various factors and OHRQoL scores.

- To compare and confirm the number of factors in OHIP-14 by using one, three, six and seven factor models available in the literature so that the discriminant validity of OHIP-14 is reassured.
- To evaluate the internal consistency and inter item correlation of OHIP-14 in measuring OHRQoL in the study population.
- To measure the responsiveness (sensitivity to change) and to classify a patient as “Improved”, “No Change” and “Not Improved”.
- To validate the findings using data from an independent cohort collected from a national sample.

2 Materials and Methods

2.1 Chapter Introduction

This chapter explains the methods used in the study described in this thesis. This includes defining the study population, sample size calculation, ethical approval for collecting data from patients and data handling. The methods of data collection are also explained. Data coding and computerisation and the various statistical methods employed to achieve the aims and objectives are discussed. The chapter also includes the statistical analysis used in this study.

2.2 Study population

The study population was comprised of people attending King's College Hospital Dental clinic (Primary Dental Care [PDC] Clinic) at Denmark Hill, London, SE5 for dental treatment. This included patients of various age groups and different ethnic backgrounds from the south of England, particularly from Camberwell and neighbouring areas. Most of the patients were referred by their dentists for treatment.

2.3 Inclusion and exclusion criteria

The patient recruitment and primary data collection were carried out by the researcher from 20th May 2013 until 4th July 2013 (Monday to Friday). All the patients who attended the King's College Dental Hospital for their dental treatment during this period who fulfilled the below mentioned criteria were approached by the researcher in order to recruit for this study. The exclusion criteria were,

- patients who were children (below the age of 18)
- patients who could not speak and understand English
- Patients who were unable to give informed consent due to mental incapacity or illness.
- People who were unable to consent themselves.

The patients who met the above criteria were excluded from this study. All other patients who were willing to take part in this study and could give consent were included.

2.4 Sample size consideration

The sample size calculation for this study involved a number of different considerations. As there were four different methodological issues namely missing data, Floor and Ceiling effects, number of factors and responsiveness of the measure to be addressed, involving four different types of analyses, the sample size was calculated for each method separately and the maximum sample required among the four methods was taken for this study in order to achieve the required power (80%) for all the analyses.

Studies involving OHIP scores mainly test the mean scores between different groups. When there is missing data in OHIP items, then the total score cannot be calculated for that particular case and hence that is excluded from the analysis. Therefore, missing data handling is adopted to increase the power of the analyses. Various possibilities involved in the sample size calculation were explored to achieve an optimum sample size in order the increase the precision of findings for this study. Slade (Slade, 2012) while reviewing the selected publications have reported an effect size of 0.4 (Gagliardi et al., 2008) in an Australian population and stated that the effect size varied between 0.2 (Crocombe et al., 2012) to 1.1 (Allen et al., 2006, Awad et al., 2000, Awad et al., 2003). Hence, both an optimal effect size of 0.4 and the lowest reported effect size of 0.2 were used for the sample size calculation. To compare the mean OHIP scores between ethnic groups (5 groups – Whites, Black, Mixed, Asian and Others) using one-way ANOVA, a total sample of 80 would be required, assuming an effect size of 0.4. A similar study would require a total sample size of 300 for the lowest effect size of 0.2. The sample size calculations are based on a two tailed test for comparing the differences with 80% power and at 5% level of significance. The sample size calculations were carried out using Gpower version 3.1.5.

To explore the floor and ceiling effects which end with a linear model to test the significant predictors of total OHIP scores, sample size calculation was carried out based on a regression model. The rule of thumb for regression equations using six or more predictors is that the ratio of number of predictors to samples should be at least 10 (VanVoorhis and Morgan, 2007). Considering the potential predictors namely age, Gender (2 categories) and Profession (with 5 categories), a total sample of 80 would suffice. However, an appropriate sample size calculation was carried out using Gpower version 3.1.5. Using total OHIP-14 scores as a dependent variable and the above three predictor variables, a study would require a total sample of 84 to find out the significant predictors assuming an effect size of 0.2 using a two tailed test at 5% level of significance with 80% power.

Different researchers have suggested different methods to calculate the sample size for carrying out factor analysis. Comrey and Lee (Comrey and Lee, 1992) suggested a sample size of 300 as good for carrying out factor analysis. The general recommendations based on the literature for carrying out factor analysis (Zhao, 2009) are summarised as below.

> Rule of 100: (Gorsuch, 1983, Kline, 1979) Gorsuch and Kline recommended at least 100 subjects (MacCallum et al., 1999). No sample should be less than 100 even though the number of variables is less than 20 (Arrindell and Van der Ende, 1985, Gorsuch, 1983).
>
> Hatcher (Hatcher, 1994) recommended that the number of subjects should be the larger of 5 times the number of variables, or 100. Even more subjects are needed when communalities are low and/or few variables load on each factor (Garson, 2008b).
>
> Rule of 150: Hutcheson and Sofroniou (Hutcheson and Sofroniou, 1999) have recommended at least 150 - 300 cases, more toward the 150 end

when there are a few highly correlated variables, as would be the case when collapsing highly multi-collinear variables (Garson, 2008a).

Rule of 200. Guilford (Guilford, 1954) suggested that *N* should be at least 200 cases in (Arrindell and Van der Ende, 1985, MacCallum et al., 1999).

Rule of 250. Cattell (Catell, 1978) claimed the minimum desirable *N* to be 250 in (MacCallum et al., 1999).

Rule of 300. There should be at least 300 cases (Norusis, 2005, Garson, 2008a).

Significance rule: Lawley and Maxwell (Lawley and Maxwell, 1971) suggested 51 more cases than the number of variables, to support chi-square testing (Garson, 2008a). According to this rule, it requires a total of 65 samples (1 case for each variable totaling 14 cases plus 51 more cases).

From all the above recommendations, the maximum sample size required is 300 and hence a sample size of 300 is sufficient to carry out factor analysis for this study.

The sample size calculation for carrying out Structural Equation Modelling suggested by MacCallum and co-workers (MacCallum et al., 1996) emphasises the use of confidence intervals for fit indices. In their work, they suggested various sample sizes for selected degrees of freedom. For this study, the degrees of freedom to carry out Structural Equation Modelling was calculated based on the formula suggested by (Rigdon, 1994) which worked out at 56 by considering 14 variables and 7 constructs. According to MacCallum and co-workers (MacCallum et al., 1999), a study with 50 degrees of freedom and 300 samples will have a power of 92.8% for close fitting and 90.3% for exact fitting. The same 300 samples will have a power of 96.0% for close fitting and 94.1% for exact fitting with 60 degrees of freedom. Therefore, this study will require a sample of 300 to make the Structural Equation Model fit the data closely or exactly with a power of more than 80%. Also, as per the popular rule of thumb of 10 cases per variable (Wolf et al., 2013), this study would require a minimum sample of 140 to carry out Structural Equation Modelling analysis.

The sample size for follow-up analysis (analysis of responsiveness) was based on the Repeated Measures ANOVA to detect the significant difference in the change in OHIP total scores over a period of three time points, within subject factors. To guarantee a power of 80% of detecting an effect size as low as 0.15, at 5% significance level, in a study with four groups and three repeated measurements, assuming a 0.8 correlation between the repeated measures, the study will require an effective sample size of 48 patients. The sample size of 48 calculated using Gpower 3.1.5 has to be adjusted for the anticipated drop outs. Anticipating a high dropout rate (more than 50% at first follow-up) at 2 months and a further (50%) at 4 months, a target sample size of 360 patients is required. Therefore, an initial sample of 360 patients was targeted for this study.

Considering all the above criteria, an optimum sample of 360 was selected for this study.

2.5 Ethical approval

Ethical approval for this study was granted on 19th March 2013 by London-Bromley Research Ethics Committee, reference number 13/LO/0366 (approval letter attached in the appendices-D and E). Similarly, the NHS R&D approval (appendix–F) was also obtained to collect data from patients at King's College Dental Hospital, Denmark Hill. The data collection was started after receiving ethical approvals.

2.6 Data collection

2.6.1 Method of Data collection

Subjects were asked to complete the questionnaire at three time points - one at presentation and at two follow-ups. The initial data collection was done at presentation when patients first attended for dental treatment at King's College Hospital Dental Clinic (Primary Dental Care). The questionnaire at presentation was completed by the patient in front of the researcher at King's College Dental Hospital when the patients came for treatment. Before starting the data collection, each participant was briefed about the objectives of this study and their role in the research project. Also, a Patient Information Sheet (PIS) about the study and a consent form were given to them by the researcher. The follow-up data were collected by sending the questionnaire by post (along with self-addressed and stamped envelope) or by emailing the questionnaire according to patients' preferences indicated at the time of baseline data collection.

2.6.2 Baseline Data

At baseline, patients were approached and explained about the study by the researcher (MA) and invited to participate. After their acceptance to take part in this study orally, they were given a study pack which contained three items namely the Information sheet about the study, a consent form and the questionnaire. The routine oral screening at King's College Dental Hospital was carried out by undergraduate dental students, regardless of this study. Patients' participation was voluntary and patients had the right to withdraw from this study at any stage without any restriction. Initially, written consent was obtained from each patient who agreed to take part in this study. Then patients were then asked to fill in the questionnaire which comprised of two parts:

1. Oral Health Impact Profile-14
2. Patients demographic details

Patients' postal address and a convenient time to contact along with their preference of contact for the follow-up were also obtained. This forms the baseline data for this research. The baseline data collection was completed on 4^{th} July 2013.

2.6.3 Data Collection Instruments

The data collection was carried out using a questionnaire which consisted of two components.

Part-1

The first part of the questionnaire consisted of the OHIP-14 instrument which contained 14 items (Slade, 1997). These items were related to their oral health and how it affected their day to day life. As an example, the first item was about the trouble they experience in pronouncing words because of their teeth/ mouth or dentures. For each item two questions were asked: one to identity whether they experience that problem, second, how much they experience it. The adopted options for the question related to whether they experience the problem included: 'Never, Hardly ever, Occasionally, Fairly often, Very often and Always'. How much the problem affected their day to day life was collected with the options: 'Not at all, Very little, Little, Much, Very much and Intolerable'. These options were included in all the 14 items in OHIP questions and the patients were asked to tick one of the options that was most relevant to them.

Apart from the 14 questions, patients were asked to rate their overall oral health by giving ratings from 0 to 10 where 0 represents their oral health was very poor and 10 represents their oral health was very good. The details about whether they had any oral health promotions programmes in the past were also collected so that the impact of this programme on OHRQoL can be compared.

Part-2

The second part consisted of demographic details of the participating individuals. This included the following:

Age: The participating individuals were asked to write their actual age in this column.

Sex: The gender of the individuals was recorded. The participants were expected to tick either one of the boxes named ‘Male’ and ‘Female’.

Ethnicity: The patients’ ethnicity was also collected. This included various ethnic groups and the participants were asked to tick the one which they belong to. The options included were: Whites (British, Irish and other), Asian or Asian British (Indian, Pakistani, Bangladeshi and others), Mixed (White and Black Caribbean, White and Black African, White and Asian and Others), Black or Black British (Caribbean, African and others) and Chinese or other ethnic group (Chinese and other ethnic background).

Relationship Status: The relationship status of the patient was also recorded. This included six options namely: Single, In a relationship, Married, Separated, Widowed and Divorced.

Education: The patient’s education level was also collected. The options included were: GCSE, A Levels, Degree (B.Sc., BA. etc.), Advanced Degree

(M.Sc., M.A., etc), Research Degree and others. If their qualification did not fit into the standard classifications given here, they were asked to fill others and write the actual qualification.

Time since last visit: Participants were asked to fill in the time of their last dental visit (prior to the current one). Various options were given under this which included ‘First time, within last 3 months, 3-6 months before, 6-12 months before, before 1 year and Don’t remember’.

Profession: The current employment status of the participants was also collected. The option included under this, were: ‘Unemployed, Full time, Part time, Student, Retired, Self-employed and others’.

Treatment Needs: The patients were asked to fill in the type of treatment they received in the current visit. The options provided under this were Periodontics, Endodontics, Orthodontics, Operatives, Surgery and Prosthetics. As majority of the patients were not aware of these technical terms used in the questionnaire, they were asked to fill in the actual treatment in their own language. Patients’ responses for this item were such as Filling, braces, root canal etc. and were later assigned a treatment category. These were later grouped into five broader categories based on their responses and the groups are Restorative, Orthodontic, Operatives/Surgery, General check-up and multiple needs. The initial grouping was carried out by the researcher with Dr. Ghotane Swapnil, a Dentist at Division of Population and Patient Health, Dental Institute, King’s

College London and then verified/ confirmed with Dr. Eduardo Bernabe, Senior Lecturer and Prof. Stephen Dunne, previous second supervisor.

Other Options such as ‘I did not want to answer the questions’, ‘I was unable to answer the question and Probably I will not continue in this study’ were included to know whether the patient was reluctant to fill in personal details or would like to drop out from the study.

Under this section, patients were asked to provide their name, mailing address, telephone or email details and the best time to contact so that the researcher could contact them for the follow-up surveys.

2.6.4 Follow-up-1

The first follow-up started two months after the baseline data collection. The questionnaires were sent in three batches and a period of one month from the date sent was considered for getting a response from the participants. The follow-up questionnaire contained only the OHIP-14 instrument (Slade, 1997). The information regarding patients’ convenient way of answering the follow-up (either e mail or postal) collected at the time of baseline data were used for the follow-up. For patients who opted for the postal questionnaire, the OHIP-14 questionnaire was posted along with a covering letter and a stamp free (postage stamp prefixed by the researcher) return envelope. Patients were asked to fill in the questionnaire and return the form using the stamp free envelope.

Patients, who opted for email response, were sent emails with an OHIP-14 questionnaire as an attachment along with a covering letter. Before sending, the OHIP-14 document was made in such a way that patients can tick the boxes for their answers (the facility available in Microsoft Word) with the help of a mouse click. Patients were asked to complete the questionnaire by ticking the answers of their choice in the electronic version and attach with the email and send back to the researcher. If no response was received within two weeks of the first email, reminders were sent to patients about the follow up if necessary and applicable. For postal questionnaires, reminders were sent as per their choice provided at the baseline data collection.

2.6.5 Follow-up-2

The second follow-up was started two months after the first follow-up. Questionnaires were sent to all the patients who took part in the baseline survey either by post or by email as per their choice given during the baseline data collection. The procedures described in the first follow-up were followed for the second follow-up also. The deadline for the last data collection was fixed at 30th November 2013. This means that participants who sent their second follow-up questionnaires on or before 30th November were considered for this study. During the first and second follow-up, in order to increase the response rate, Dillman's approach was adopted (Dillman, 1978). According to this, the non-response rate can be significantly reduced by adopting mixed mode surveys and survey techniques such as mailed questionnaire, electronic means, telephone, follow-up reminders, anonymity of response and return postage.

Patients who had given permission to be contacted by phone were reminded about the follow-up questionnaire through a phone call.

2.7 Data Handling

2.7.1 Data checking and cleaning

The self-reported data were anonymised by giving a unique identification number (ID) for each participant and the personal details were securely stored. The follow-up questionnaires were given the same ID numbers as baseline data. The baseline and the follow-up data collected using paper questionnaires were stored in a locked cupboard which was accessible to only the researcher and the supervisors. The follow-up questionnaires collected through emails were stored in a password protected computer kept inside a locked room. Both paper based and electronic data were kept at the Biostatistics and Research Methods Centre, King's College London Dental Institute, Denmark Hill campus and were accessible only to the researcher and the supervisors. Once the study is completed, the completed questionnaires will be destroyed and the electronic questionnaires will be deleted from the computer.

2.7.2 Data coding and computerisation

The collected data in the text form were then quantified using appropriate coding schemes to facilitate quantitative analysis. The data were classified as nominal, ordinal and quantitative (a continuous) data and coding was carried out accordingly. After computerising the data, the coded data were checked manually by comparing the data in the questionnaire with the computerised data to avoid any discrepancies or errors in the process of computerisation. Also, run control checks such as checking the range for variables, column totals were carried out in the process for further cleaning. The coding scheme used in this study is given in the appendix G.

Each participant was given an ID number so that they could be identified in future and the ID number started from 1 to 360. The numbers were followed by A or B or C representing Baseline, Follow-up-1 and Follow-up-2 respectively. Before that, a database in SPSS was created by defining variables along with their properties to enter the data. After coding the data for each participant, they were entered into the SPSS database. Again, the electronic data were given two levels security as 1. Locked room (Physical security) and 2. Password protected computer (Electronic security). The personal details were not exposed to any one apart from the researcher and the supervisors.

The responses for the items in some of the demographic variables were too low to be able to make meaningful comparisons. Therefore, such items of some of the demographic variables were combined for the analysis as below.

Ethnicity

- British, Irish and Other whites were combined and named as ‘Whites’.
- Indian, Pakistani, Bangladeshi and other Asians were combined and named as ‘Asian’

- Caribbean, African and Other blacks were combined and named as 'Black'.
- White and Black Caribbean, White and Black African, White and Asian and Other mixed cases were combined and named as 'Mixed'.
- Chinese and other ethnic background were combined to form 'Others'.

Education

Degree (BA, BSc, etc), Research Degree and Advanced degree (MA, MSc, etc) were combined and named as Degree and Higher.

Profession

Retired, OAP and Pensioner were combined as 'Retired'.

'Others' Category included the following.

Actress, Cleaning, Disability benefits, Disability sick, Home maker, Housewife, Painter, Sculptor, Supply teacher, Volunteer and others.

Treatment Needs

The responses for this item were grouped in to five categories as below. The details of the groupings are explained under section 4.6.3.

1. Restorative

   Periodontics, Endodontics, Prosthetics (root canal, bridge, Crown, Partial denture, tooth grinding, false teeth, fillings, checking Gums and plates).

2. Orthodontics

3. Operatives/ Surgery (Operatives, Surgery, implants and extraction)

4. General Check-up (check-up, General treatment and infection)

5. Multiple needs (Patients who required more than one of the above categories)

2.8 Statistical Analysis

The statistical analyses used to address various objectives are detailed below.

Objective 1

- Descriptive statistics including the mean, standard deviation, median, minimum, maximum and percentages were used to summarise the sample characteristics (Demographic details).
- Summary statistics for the total OHIP scores (composite score) were calculated for each dimension.

Objective 2

The missing data pattern in OHIP items was explored and the missing data were filled with the following missing data handling techniques:

- completed case analysis
- filled with item mean
- filled with subject mean
- With interpolation
- Regression models with linear trend
- Estimation Maximisation Algorithm
- multiple imputation

The mean OHIP scores and standard errors for each of the methods were compared.

Objective 3

- the missing data pattern in the follow-up data was explored using logistic regression and graphical methods.
- The OHIP-14 scores were taken at baseline, 2 months and 4 months after treatment. The missing values within the OHIP items were filled in by the suitable method identified from Question-1.
- As there was a high percentage of missing data in the total OHIP-14 scores, the data were analysed and compared using multilevel modelling (Random Intercept Model) which includes all the available data in the analysis to get improved estimates

for the predictor variables. It assumes that error terms at every level of the model are normally distributed and equality of population variances (homoscedasticity) (Rabe-Hesketh and Skrondal, 2012).

Objective 4

Floor and ceiling effects were quantified and the Tobit regression method to handle these effects by adjusting these effects in the model, was used to explore the actual relationship between the OHIP composite scores and other factors.

Objective 5

To confirm the number of factors in OHIP-14, four different models (one factor, three factor, six factor and seven factor models) were considered. The suitability of these models for the study data were checked using Structural Equation Modelling. The fit indices were then compared for all the four models.

Objective 6

The internal consistency of the items in OHIP-14 were measured using Cronbach’s alpha. The correlation between the items were measured using inter-item correlation coefficient.

Objective 7

The responsiveness of OHIP-14 to a change in Oral Health Related Quality of Life was studied using the method suggested by Oord and Ark (Oord and Ark, 1997) and the patients were classified as “Improved”, “No change” and “Not Improved” based on the suggested measure .

Objective 8

The findings of the above objectives were tested using the national data from Adult Dental Health Survey 2009,

3 Description of Sample and OHIP scores

3.1 Chapter introduction

This chapter addresses the objective of Quantifying the OHRQoL using the study data by describing the sample characteristics, analyse OHIP scores based on completed cases analyses in order to find out the significant predictors of total OHIP and to examine the psychometric properties (Internal consistencey and inter item correlation) of OHIP-14. In the conventional way, the data were analysed using completed cases analysis, meaning that subjects who had data in all the 14 items were included in the analysis. Where one or more OHIP items were missing, those cases were excluded from the analysis. The commonly used statistical software packages such as SPSS and Stata implement this as a default setting. As there were 277 respondents who had filled in all the 14 items (the remaining 83 had one or more items missing) at baseline, these analyses were restricted to these 277 cases only. The remaining 83 were excluded from the analysis. The analysis included descriptive statistics of the sample characteristics and OHIP scores.

3.2 Descriptive and demographic details

A total of 360 participants took part in this study, out of which 277 provided data on all the 14 items and the remaining 83 had one or more missing items. However, the demographic details were available for most of the patients who took part in this study. The sample comprised 148 (41.11%) males and 212 (58.89%) females. The participants were in the age range of 19 to 84 years and the mean age of the participants was 45.76 years with a standard deviation of 15.10 years. As far as dental visits were concerned, 78.61% had visited a dentist within the past 6 months and 3.06% were first timers. Eight of the participants did not remember their last dental visit. The other demographic details of the samples are shown in Table 3.1.

Table 3.1 Demographic details of the study sample

Variables	N (%)
Gender	
Male	148 (41.11)
Female	212 (58.89)
Age	
Mean (sd)	45.76 (15.10)
Min, Max	19, 84
Ethnicity	
Whites	246 (68.33)
Black	59 (16.39)
Mixed	30 (8.33)
Asian	17 (4.72)
Chinese and Others	4 (1.11)
Missing	4 (1.11)
Relationship	
Single	110 (30.56)
In a relationship	81 (22.50)
Married	110 (30.56)
Separated	16 (4.44)
Widowed	7 (1.94)
Divorced	24 (6.67)
Missing	12 (3.33)
Education	
GCSE	77 (21.39)
A Levels	50 (13.89)
Degree or higher	161 (44.72)
Others	39 (10.83)
Missing	33 (9.17)
Employment	
Unemployed	40 (11.11)
Full time employed	167 (46.39)
Part time employed	48 (13.33)
Student	15 (4.17)
Retired	49 (13.61)
Others	32 (8.89)
Missing	9 (2.50)
Last Dental Visit	
First time	11 (3.06)
within 3 months	194 (53.89)
3-6 months before	89 (24.72)
6-12 months before	31 (8.61)
longer than 1 year	14 (3.89)
Don't remember	8 (2.22)
missing	13 (3.61)

3.3 Description of Oral Health Related Quality of Life for completed cases

The mean total OHIP-14 score for the samples was 22.81 (sd=16.10) with the median score of 20.00 (range 0 to 61). Out of the maximum possible composite (total) score of 70, Females (24.17) had slightly higher mean score than males (20.97). However, the difference was not statistically significant ($p=0.14$). The median total score for each ethnic group is shown in Figure 3.1. The total score did not differ significantly between different ethnic groups ($p=0.99$).

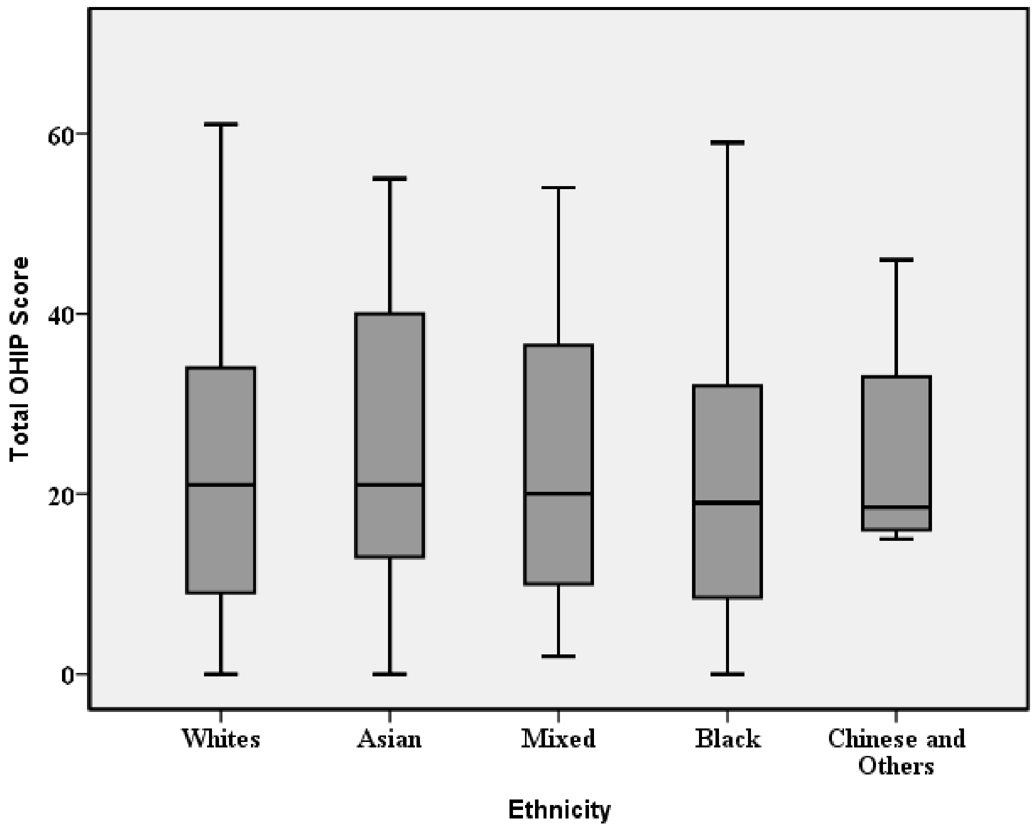

Figure 3.1 Box plot depicting the median OHIP-14 total score for different ethnic groups.

Note: The box represents the Inter Quartile Range (IQR) (the difference between first and third quartile), the middle line represents the median and the two whiskers represent -1.5*IQR and +1.5*IQR.

3.3.1 Predictors of Oral Health Related Quality of Life based on completed cases

The potential predictors of OHIP total score namely age, gender, relationship, profession, last visit to dentist, Education and Ethnicity were included in the multivariate regression model and the total OHIP score was the dependent variable. The results of multivariate regression analysis are shown in Table 3.2. The results showed that Relationship, Profession, Last Dental Visit and Education significantly predicted total OHIP scores. When compared to ‘Full time’ group, the ‘Unemployed’ group had significantly higher (p=0.03) total OHIP scores indicating a worse OHRQoL. Similarly, in education, when compared to ‘Degree or higher’ group, patients who had ‘Other’ qualifications such as Diploma, had a significantly higher (p=0.03) OHIP total score indicating poor oral health related quality of life. When compared to people who remained "Single", People who were "Separated" had significantly higher (p=0.04) total OHIP scores and hence poorer Oral Health Related Quality of Life. People who had visited dentists more than a year ago showed significantly lower (p=0.04) total scores when compared to people who had visited a dentist within 3 months.

Table 3.2 Results of multivariate regression for predicting total OHIP score

Predictor	Comparison group	effect	95% CI		P value
			LCL	UCL	
Age		-0.04	-0.23	0.14	0.63
Gender	Male				
Female		3.49	-0.79	7.77	0.11
Relationship	Single				
In a Relationship		3.12	-2.58	8.84	0.28
Married		1.66	-4.08	7.39	0.57
Separated		11.25	0.65	21.85	0.04
Widowed		3.75	-15.16	22.66	0.70
Divorced		7.71	-0.85	16.27	0.08
Profession	Full time				
Unemployed		7.90	0.90	14.90	0.03
Part time		0.90	-6.00	7.80	0.80
Student		0.20	-11.16	11.55	0.97
Retired		-7.63	-15.60	0.34	0.06
Other		-1.02	-8.22	6.17	0.78
Most recent dental visit	< 3 months				
First time		-0.02	-11.77	11.72	0.99
3-6 months		0.26	-4.68	5.20	0.92
6-12 months		-8.88	-17.33	-0.43	0.04
More than 1 yr		1.29	-11.03	13.62	0.84
Don't remember		7.75	-5.98	21.48	0.27
Education	Degree or higher				
GCSE		4.42	-0.70	9.55	0.09
A level		-2.61	-8.85	3.63	0.41
Other		7.64	0.68	14.60	0.03
Ethnicity	Whites				
Asian		3.12	-6.06	12.30	0.50
Mixed		-1.08	-9.26	7.09	0.79
Black		-1.35	-7.16	4.45	0.65
Chinese and Others		3.22	-13.19	19.63	0.70

3.3.2 OHRQoL by seven dimensions

The Oral Health Related Quality of Life scores were analysed according to seven dimension defined by Slade (Slade, 1997). Table 3.3 summarises the mean (sd) and median (min, max) OHIP scores for the seven dimensions. The maximum mean score of 4.37 was obtained for psychological discomfort, indicating that dental problems affected the patients psychologically. This was followed by physical pain which had a mean score of 4.34 with standard deviation of 2.72. These two dimensions scored more of less the same score indicating that psychological discomfort and physical pain were the foremost problems people faced due to their dental conditions in this cohort. The lowest mean score of 1.60 with a standard deviation of 2.24 was observed for functional limitation. A higher score of 3.87 (sd=2.82) was observed for psychological disability and the remaining two dimensions of social disability and handicap had similar scores (2.77 and 2.49 respectively). As the overall OHIP-14 score was slightly higher for females when compared to males, a further subgroup analysis including gender and dimension revealed that the mean score for the dimension, Psychological discomfort was significantly higher for females, mean(sd)= 4.72 (2.93) than males, 3.88 (2.97). However, no statistical significance was obtained for the other dimensions.

Table 3.3 Dimension-wise summary of OHIP scores

Dimension	mean(sd)	median (min, max)
Functional Limitation	1.60 (2.24)	0 (0, 10)
Male	1.63 (2.29)	1 (0, 10)
Female	1.59 (2.21)	0 (0, 10)
Physical Pain	4.34(2.72)	4 (0, 10)
Male	4.15 (2.63)	4 (0, 10)
Female	4.48 (2.77)	5 (0, 10)
Psychological discomfort	4.37 (2.97)	4 (0, 10)
Male	3.88 (2.97)	4 (0, 10)
Female	4.72 (2.93)	5 (0, 10)
Physical disability	2.64 (2.67)	2 (0, 10)
Male	2.42 (2.56)	2 (0, 9)
Female	2.80 (2.74)	2 (0, 10)
Psychological disability	3.87 (2.82)	4 (0, 10)
Male	3.56 (2.83)	3 (0, 10)
Female	4.10 (2.81)	4 (0, 10)
Social disability	2.57 (2.77)	2 (0, 10)
Male	2.31 (2.70)	1 (0, 10)
Female	2.74 (2.80)	2 (0, 10)
Handicap	2.49 (2.78)	2 (0, 10)
Male	2.44 (2.83)	2 (0, 10)
Female	2.52 (2.75)	2 (0, 10)

The patients were grouped according to their treatment needs. As only 218 participants provided information on their treatment needs, out of which only 144 had completed all the items in OHIP-14, this analysis was based on these cases only. Figure 3.2 shows the mean OHIP score along with 95% confidence intervals according to treatment needs. Though subjects with multiple needs had a higher mean total OHIP score, no significant differences were seen between subjects with different treatment needs by one-way analysis of variance (p = 0.13). Similarly, there were no significant differences between subjects in different occupations (p = 0.10) (Figure 3.3).

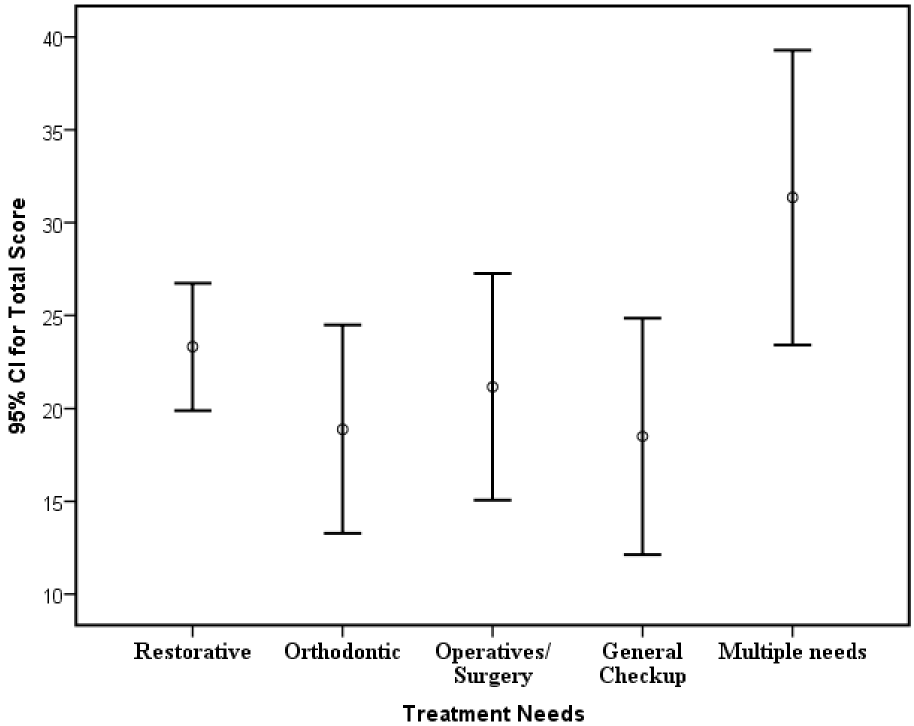

Figure 3.2 The mean and 95% Confidence interval of total OHIP scores for various treatment needs

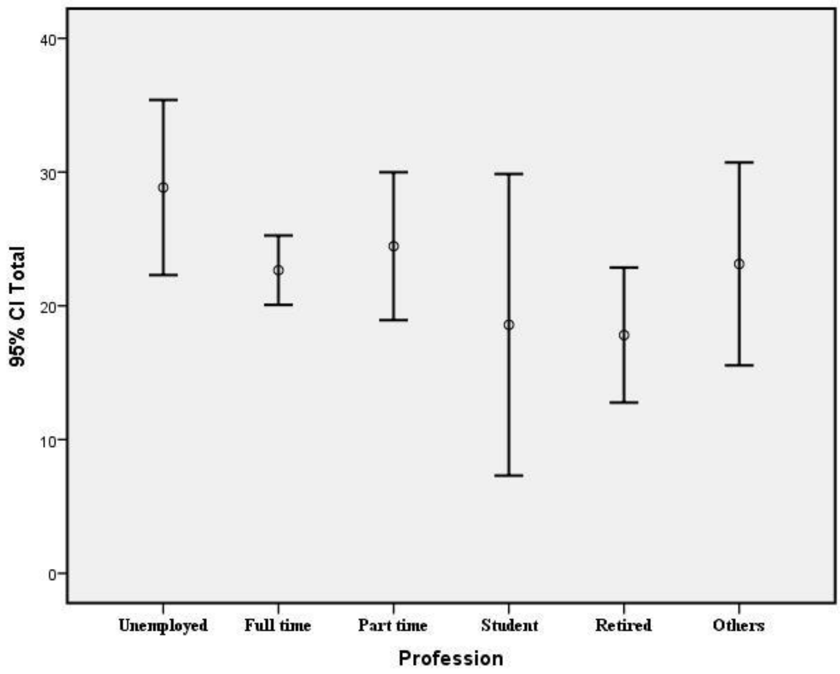

Figure 3.3 The mean and 95% Confidence interval of total OHIP scores for various Professions.

3.4 Psychometric properties of OHIP-14

Internal consistency of the data was assessed using Cronbach's alpha, where a score higher than 0.60 is considered good to excellent reliability (Locker and Slade, 1993). The Cronbach's alpha based on this data was found to be 0.94, indicating that the items in this instrument are internally consistent in measuring the Oral Health Related Quality of Life (OHRQoL). The correlations between the items are described for all the 14 items in Table 3.4.

Table 3.4 Bivariate correlation between different OHIP-14 items

Items	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14
Q1	1.0	0.49	0.24	0.42	0.43	0.34	0.36	0.36	0.40	0.45	0.38	0.43	0.43	0.36
Q2		1.00	0.32	0.44	0.42	0.37	0.41	0.44	0.40	0.42	0.37	0.39	0.46	0.37
Q3			1.00	0.56	0.37	0.47	0.43	0.53	0.60	0.29	0.52	0.47	0.45	0.51
Q4				1.00	0.50	0.61	0.60	0.71	0.68	0.45	0.52	0.53	0.52	0.49
Q5					1.00	0.72	0.46	0.52	0.57	0.84	0.56	0.49	0.63	0.42
Q6						1.00	0.59	0.64	0.74	0.69	0.68	0.61	0.70	0.54
Q7							1.00	0.72	0.66	0.52	0.58	0.59	0.63	0.48
Q8								1.00	0.72	0.50	0.59	0.61	0.58	0.52
Q9									1.00	0.57	0.69	0.66	0.67	0.61
Q10										1.00	0.60	0.52	0.65	0.41
Q11											1.00	0.72	0.69	0.58
Q12												1.00	0.66	0.69
Q13													1.00	0.58
Q14														1.00

A maximum correlation of 0.84 was observed between items 5 and 10 and the minimum correlation of 0.24 was observed between items 1 and 3. As the correlation between the items were mostly above 0.40, it is clear that the items are highly correlated. However, none of the items had a correlation of 0.85 and above indicating no duplicate items in the OHIP-14.

Table 3.5 Item total statistics showing the relationship and the impact of each item in OHIP-14 instrument.

Question No	Scale mean if item deleted	Scale variance if item deleted	Corrected item-total correlation	Squared multiple correlation	Cronbach's alpha if item deleted
Q1	21.97	236.67	0.51	0.35	0.94
Q2	21.86	236.29	0.51	0.35	0.94
Q3	20.67	229.47	0.59	0.45	0.94
Q4	20.55	224.47	0.72	0.61	0.94
Q5	20.35	220.31	0.74	0.77	0.94
Q6	20.77	219.56	0.81	0.76	0.94
Q7	21.65	226.95	0.73	0.63	0.94
Q8	21.18	222.49	0.77	0.71	0.94
Q9	20.93	217.94	0.85	0.78	0.93
Q10	20.71	221.27	0.73	0.76	0.94
Q11	21.29	221.57	0.77	0.65	0.94
Q12	21.64	223.74	0.76	0.69	0.94
Q13	21.16	218.21	0.80	0.67	0.94
Q14	21.76	225.28	0.68	0.61	0.94

Table 3.5 depicts the various measures such as scale mean, variance and Cronbach's alpha if an item is deleted along with corrected item-total correlation and squared multiple correlation to explore the internal consistency of the items.

The scale mean and variance if item deleted show the mean and variance if that particular item is being removed. Among the 14 items, none of the items showed a major decrease or increase in mean or variance of the total score when that particular item was removed, which indicates that all the 14 items are internally consistent. The corrected item total correlation depicts the Pearson correlation coefficient between the item and the total score excluding the item. The higher correlation indicates the better relevance of the item. The correlation values were relatively high and the maximum corrected item total correlation was 0.85; none of the items had a correlation value less than 0.50. The squared multiple correlation, is the R^2 value of an item when it is predicted from all other items in the scale. Dropping of an item is to be considered if the R^2 value is less than this (preferably if less than 0.30). However, in this case, the lowest value was 0.35 and all the items have reasonably good R^2 indicating that all the items contribute to the internal consistency of the instrument. The “Cronbach’s alpha if item was deleted”, is the estimated value of alpha if the given item was removed from the model. If this value is higher than the overall alpha value then that particular item may be deleted in order to improve the overall alpha value. But in this case, there was no change in alpha value and all the values were close to 0.93 and 0.94 for all the items indicating the importance of all the 14 items in OHIP-14.

3.5 Discussion and Conclusions

The descriptive analyses showed that there was no significant gender difference with respect to total OHIP scores, which is in keeping with findings of

Gisler and Stenman (Gisler et al., 2012, Stenman et al., 2012). The high Cronbach's alpha and good inter-item correlations showed that OHIP-14 has good psychometric properties. Further sub-group analysis showed that women experienced more psychological discomfort than men. An interesting finding is that the ethnic groups did not differ with respect to the oral health related quality of life. All ethnic groups experienced similar OHRQoL of life which may be of importance for intervention planning. The dimension wise analysis showed that physical pain and psychological discomfort are the main factors that determined the OHRQoL. Separate analysis based on 'treatment needs' using one-way ANOVA showed no significant difference between different 'needs' groups with respect to the composite scores, indicating that the type of treatment or the problem did not influence the OHRQoL.

The significant predictors of OHIP total scores are relationships, profession, last dental visit and education. People with full time employment experience better OHRQoL when compared to unemployed people. A negative relationship between community level unemployment and population's preventive oral health care utilization has been reported by (Quinn et al., 2009). Also, they have concluded that Community-level unemployment may impede or distract populations from utilizing preventive dental services. People with other qualifications such as Diploma had significantly higher total scores compared to people with Degree or higher qualification. Though the multiple regression analysis based on completed cases provided these factors (relationships, profession, last dental visit and education) as significant predictors of OHIP-14 total scores, it is difficult to find a pattern to arrive at a conclusion for any one of the significant predictors. As an example, for the relationship status, ‘Separated’ was one of the significant predictors whereas ‘Divorced’ is not.

The results obtained from the analyses described in this chapter, were carried out on the completed cases only, resulting in the exclusion of 83 subjects’ data. Depending on the distributions of these missing data points, this might result in bias in the sample and will lead to loss of Power for the study. Therefore, these results were compared with later analyses based on the procedures to handle missing data in the dataset.

4 Missing Data Analysis

4.1 Introduction

The objectives of exploring missing data mechanism, comparison of various missing data methods, evaluating the bias introduced by the missing data in the longitudinal studies and the suitable method to treat the bias were addressed in this chapter. Handling missing data from experiments is of considerable concern to researchers, specifically in studies which are questionnaire based, as missing data may have important impacts on the results, particularly when the initial sample size is small. As previously discussed, in oral health research, the effect of dental treatment on quality of life of individuals are often investigated using Oral Health Related Quality of Life (OHRQoL) assessment. As the patients themselves fill in the questionnaire, the problem of missing data is very commonly encountered in such assessments (Gomes et al., 2016). Despite this, there are very few studies (Allen et al., 2009, Awad et al., 2014, Baba et al., 2008, Gisler et al., 2012, Reissmann et al., 2013, Royston, 2004, Saub et al., 2005) that have reported on the presence of missing data and described the methods used to handle such an issue. In particular, no attempts appear to have been made to compare the use of different techniques for managing missing data in terms of efficiency in handling missing data in the field of dental research.

Missing Data Mechanism

The missing data in any dataset follows one of the following patterns:

1. Missing Completely at Random (MCAR)
2. Missing At Random (MAR)
3. Missing Not At Random (MNAR).

Missing Completely At Random (MCAR)

In this pattern, the missing data in the sample are completely unrelated to the values of the variables (both observed and missing values) included in the study. As an example, in laboratory studies, if a researcher dropped a sample by mistake, then the missing data due to dropping belongs to this mechanism.

Missing At Random (MAR)

The missing data that follows this pattern is unrelated to any of the missing values but may be related to the values of other study variables. As an example, missing income data may be related to social class as the higher social class may be reluctant to give the income but not related directly to actual income values. If the missing data follows MAR pattern or MCAR pattern then a number of different techniques are available to handle this situation.

Missing Not At Random (MNAR)

With this pattern, managing the missing data is more challenging and cannot be ignored. In this case, the missingness occurs not by random and the missing value pattern depends on the missing values. As an example, when individuals are not providing their age purposely as they feel they are old, then such a value belongs to this category.

Though missing data poses a serious concern, the majority of statistical packages do not include those cases even if it has only one item missing. So, if the size of missing data is large, (it is recommended that 5% or less of missing data are insignificant (Schafer, 1999), the actual samples included in the analysis will be small and it reduces the power of the study to detect differences. However, other researchers have stated that statistical analysis without considering missing data will be biased if the missing data are 10% or more (Bennett, 2001). In the case of large samples, though the impact of missing data may be minimal, the pattern of missing data has to be explored before carrying out further analyses as it affects the findings. When compared to percentage of missing data, missing data patterns and the mechanism have more impact on the results (Tabachnick and Fidell, 2013).

The sources of missing data can be many fold and in OHIP, two types of missingness – item missing and case missing, which need to be tackled.

This chapter aims to identify the patterns of missing data in both item missing and in case missing in the follow-up data and to explore suitable methods to handle these occurrences.

4.2 Methods for dealing with Missing Data

4.2.1 Listwise Deletion

It is common practice in medical and dental research to handle missing data by listwise deletion (Peugh and Enders, 2004). This is where a subject's entire record is omitted from analysis if any single value is missing (Completed case analysis). Although this method is unbiased under the assumption of missing data that categorised as MCAR, it leads to incorrect estimation and error in results when the missing data are classified as MAR or MNAR (Osborne, 2012). As the missing data cases are deleted, this leads to smaller samples for the analyses which affect the power of the study. But this method is undesirable and often leave the null hypothesis unchallenged when the sample size is small because larger sample size results in greater ability to reject the null hypothesis correctly in favour of alternative hypothesis (Dupont, 2014). Thus, the listwise deletion method may be suitable to handle missing data when the study involves a reasonably large sample, the percentage of missing data is small (less than 5%) and the missing data mechanism follows MCAR. However, it is problematic from the perspective of power in smaller samples.

4.2.2 Mean substitution

Another classic method of handling missing data is to substitute the missing data with subject mean or with item mean. Though these two methods provide similar results to other methods, it is not appropriate to use these data in further analyses as these methods do not look into the pattern of missingness. This method is applied under the assumption that when no other information is available, the mean is the better estimate of the missing values. But when the percentage of missing data is relatively large, substituting them with a mean score reduces the variance of the variable and hence the effect of missing data on the results will also increase. The practice of mean substitution in creating the composite scores in multiple items of a questionnaire may alternatively be justified on condition that the internal consistency of the instrument used is relatively high (0.90 or more) and the inter item correlation is high (0.70 or more) (Osborne, 2012).

4.2.3 Interpolation, Trend and Regression methods

Interpolation, trend and regression methods have significant advantages when compared to use of mean substitution. These methods use equations to estimate / predict the missing values. As the estimation is based on the available values and adopts statistical methods of estimation like maximum likelihood, the estimators are unbiased. These methods are somewhat better than the mean substitution as these methods look into the relationship between the variables for substituting the missing data. These methods are considered to be single imputation as they are concentrating on substituting the missing value with a single measure. The single imputation is the method where the missing values are estimated using regression models or some other interpolation methods and the estimated values are substituted once only. No further estimation is carried out.

4.2.4 Expectation Maximization Algorithm (EM Algorithm)

This method involves two steps, namely an 'E' step and an 'M' step. In the 'E' step, the conditional expectation of the missing data is found using the observed values and the current estimates of the parameters. These expected values are then substituted for the missing data and the 'M' step estimates the maximum likelihood estimate of the parameters. As this method uses an iterative process and the likelihood procedure assuming a normal distribution to estimate the missing data, it is considered to be one of the more reliable methods used for handling missing data (Pigott, 2001).

4.2.5 Multiple Imputation (MI)

In recent times, the method of Multiple Imputation has become commonly used (Sterne et al., 2009) among statisticians due to the advancement in computing power. In this method of imputation, the uncertainty due to missing data is handled by generating ‘m’ plausible values for each missing data point and hence it generates ‘m’ data sets. Then these ‘m’ data sets are analysed separately and the results of individual data sets are then pooled to get a parameter estimate which incorporates the uncertainty due to missing data (Dong and Peng, 2013). The estimates obtained using multiple imputation are less biased (Graham et al., 2007) and closer to the original values. As the number of imputations increases, the difference between the original and estimated values reduces. In most of the realistic circumstances, a total of 2 to 10 imputations will suffice (Donald, 1987, Little and Rubin, 2014). The processes involved in multiple imputations are depicted in Figure 4.1.

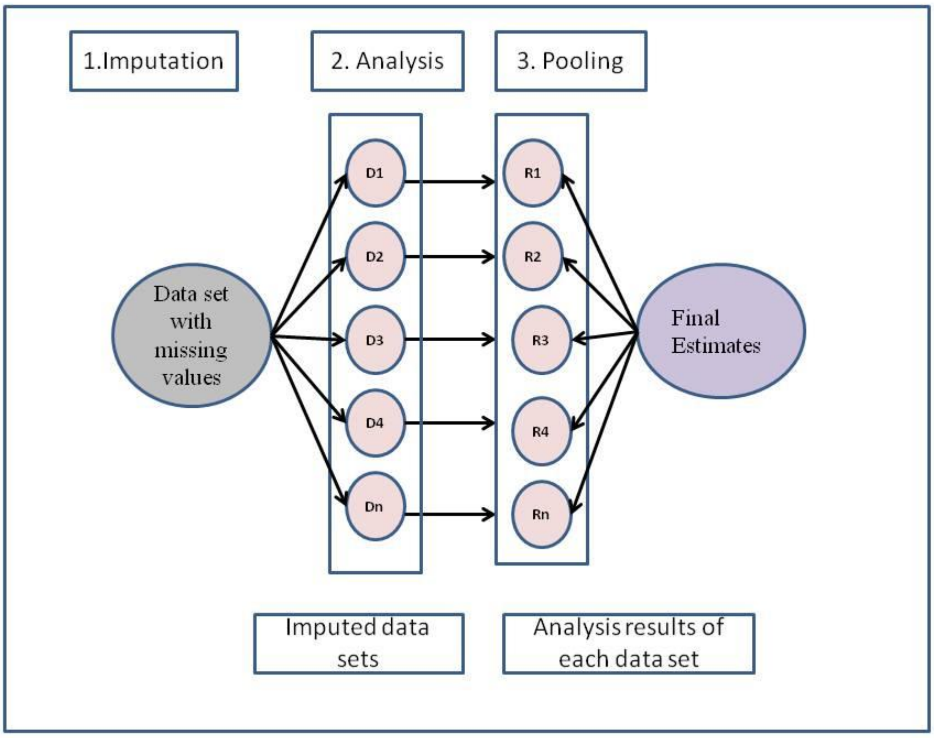

Figure 4.1 Processes involved in Multiple Imputation

In the above process, the missing values in the data set are filled with imputed values where the imputation is based on suitable methods such as regression, interpolation etc. Each time it imputes, a new data set with imputed values are created. Usually, ten imputations are considered reasonable, however, the number of imputations depends on the researcher and the level of accuracy required. The imputed data sets are represented as D1, D2, .. Dn in the above diagram. Each imputed data set is analysed separately and results are stored (represented as R1, R2, … Rn in the above diagram). Then the results from each data set are pooled to get a single figure as final estimates.

4.3 Study Methods

As there are several methods available to handle missing data as discussed above, the selection of the appropriate method is critical for achieving an unbiased result. The missing data in OHIP-14 items could be due to either the participants failing to respond to one or more items or not paying attention to the question. However, these individuals will potentially provide good data on other items to contribute to the overall analysis based on the total scores. When handling missing data in the individual items of OHIP-14, in order to get the total score at a particular time point (cross sectional) there is no question of an individual being excluded completely from the study (data missing in all items) because the interest here is to create a composite score based on many items in the instrument and all the patients included in the study that were recruited at baseline.

The missing data points in the OHIP-14 items were filled with the above methods and the summary statistics were calculated separately for each method. The suitability of these methods was considered by comparing the summary statistics obtained from these different methods.

4.3.1 Missing data at baseline

At baseline, the missing items were explored to find out the pattern as discussed before. Each item with the number of missing cases along with percentages was presented to have an overall idea about missing items at baseline. A pie chart was used to depict the percentage missing values, variables and items separately. Missing value patterns for each item were also presented diagrammatically. Little's MCAR (Little, 1988) test was used to find out whether the missing items were Missing Completely At Random or not.

4.3.2 Missing Data Pattern in the Follow-up data

The summary statistics were used to depict the total OHIP scores for demographic variables at various follow-up time points. The missingness of the total score was explored separately for the different follow-up periods. Hence, to assess whether subject characteristics such as age, gender and ethnic group predicted dropout, a logistic regression model was used to find out the effect of baseline subject characteristics on the dropout. Also, to find out whether the earlier total scores predicted the later dropouts, participants at time t (specific time point) were grouped as non-dropouts if they provided data at time t and also at time t+1 (the next time point) (non-dropout group). Similarly, the participants at time t were grouped as dropouts if they supplied data at time t and failed to do so at time t+1. For these two groups (dropout and non-dropout), the means and 95% confidence intervals of total scores were then plotted to assess the emerging pattern.

As there was no evidence that the earlier values predicted dropouts from Figure 4.4, it may be assumed that the dropout mechanism in total scores over the follow-up period happened completely at random (MCAR). However, it is necessary to explore if any of the demographic variables predict the dropouts.

Using the logistic regression analysis (Table 4.2), it was observed that both age and gender predicted the dropouts, indicating that observed demographic variables did predict the dropouts. Hence, Random Effects Models were used for the follow-up analysis to find out the predictors of OHIP scores as the inferences drawn from such models remain valid under the less restrictive assumption of Missing At Random (MAR). This allows the dropout to be either predicted by earlier values or by the covariates.

For the total scores at various time points, the Random Intercept model was fitted using the “xtmixed” command in Stata 12.0. In the Random Intercept models, the varying intercept accounted for the correlation between the repeated total scores per subject and the fixed part of the model (regression) contained the contrasts for the time factors (baseline, 2 months after and 4 months after), the main effects of gender, sex and the interactions. The various interaction effects were tested for significance and the interaction terms were included in the final model only if they were significant. Otherwise, the final model included only the main effects. If evidence for interaction was found, then further post hoc analysis were carried out to assess the temporal changes. If so, such post hoc comparison was adjusted for multiple testing using Bonferroni corrections.

4.4 Results

4.4.1 Exploring missing OHIP items in baseline data

At baseline, a total of 83 (23.06%) participants had one or more items missing and the remaining 277 provided complete data on all the 14 items. Among the 83 participants who had one or more missing items, 38 participants had one item missing and two items were missing for 18 of the participants. Between 3 and 7 items were missing for 26 of the participants whereas only one participant had 11 items missing. The item wise missing values in the OHIP-14 are given in Table 4.1. The percentage of missing items ranged between 1.39% and 7.50%. A maximum of 7.50% missing items was observed for item 4 which asked whether they found it uncomfortable to eat any foods because of dental problems.

Table 4.1 Item wise percentage of missing data.

Item No.	Completed cases	Cases with missing items	%
1	355	5	1.39
2	350	10	2.78
3	339	21	5.83
4	333	27	7.50
5	346	14	3.89
6	347	13	3.61
7	338	22	6.11
8	344	16	4.44
9	343	17	4.72
10	344	16	4.44
11	346	14	3.89
12	348	12	3.33
13	345	15	4.17
14	347	13	3.61

A minimum of 1.39% missing items was noted for item 1 which deals with problems in pronouncing words because of the dental problems. The percentage of missing data in the items affect the total OHIP score for the participants and hence any analysis without considering the missing data / the mechanism may lead to bias in the outcomes. As the percentage of missing data was more than 5.00% for some of the items, it was important to explore the missing data mechanism. The details of missing data in terms of missing variables, missing cases and missing values are given in Figure 4.2.

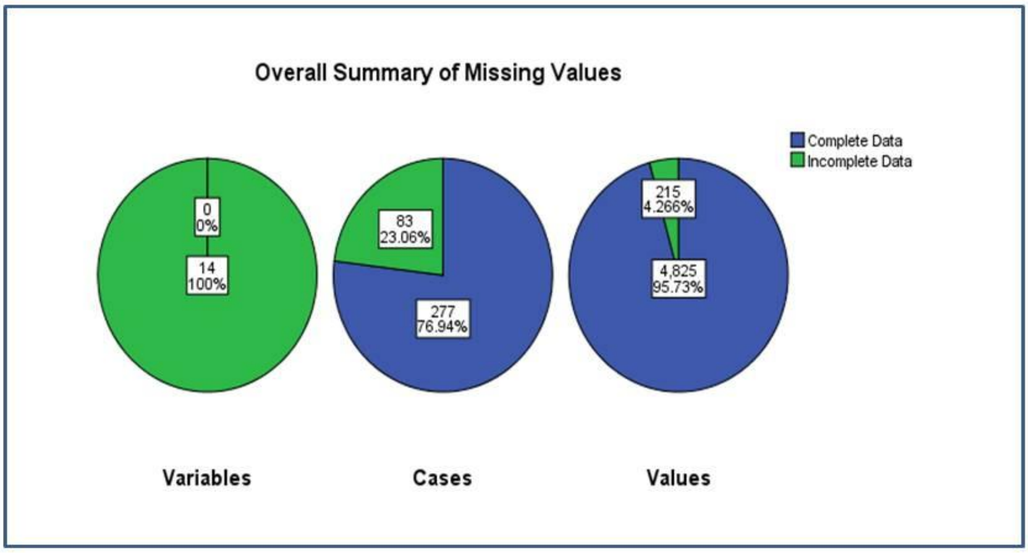

Figure 4.2 Missing variables, cases and values in OHIP items

The data were analysed to find out whether the missing data in the items follow any one of the missing patterns described above. Little's MCAR test (χ^2=618.42, p=0.16) showed that the missing data in this study followed the Missing Completely At Random (MCAR) pattern. To find out whether any of the demographic variables predicted the missingness, a logistic regression analysis was carried out using missingness as outcome measure and age, sex, ethnicity, education and profession as predictor variables and the results are shown in Table 4.2. This showed that none of the variables predicted ($p>0.05$ for all variables) missingness. Hence, it supports the conclusion that the missing data in OHIP-14 items followed the Missing Completely At Random pattern.

Table 4.2 Results of Logistic Regression to find out the predictors of missingness

Variable	Comparison Group	Effect	Odds Ratio	95% CI		P value
				LCL	UCL	
Age		0.012	1.01	0.99	1.04	0.29
Gender	Male	0.38	1.47	0.82	2.62	1.47
Ethnicity	Whites					0.77
Asian		-0.25	0.78	0.21	2.93	0.71
Mixed		0.54	1.72	0.71	4.15	0.23
Black		-0.06	0.94	0.43	2.04	0.88
Other		-19.85	0.00	-	-	0.99
Education	GCSE					0.32
A level		0.72	2.05	0.80	5.23	0.13
Degree		0.69	1.99	0.93	4.27	0.08
Other		0.67	1.95	0.72	5.25	0.19
Profession	Unemployed					0.31
Full-time		0.56	1.75	0.60	5.07	0.31
Part-time		0.62	1.87	0.55	6.33	0.32
Student		0.38	1.47	0.23	9.43	0.69
Retired		0.33	1.39	0.36	5.38	0.64
Other		0.91	2.50	0.70	8.90	0.16

'-' values could not be estimated due to small samples in that particular group.

The various missing data patterns are shown in Figure 4.3. The chart displays various missing data patterns found in the data set. The cases with the same pattern of incomplete and complete data are represented in each column. As an example, pattern 1 represents the cases that have data for all the 14 items (no missing values). Pattern 11 has all cases that have missing data on question 6 and question 11 only. Similarly, Pattern 56 includes all the cases that have missing data on questions 1, 3, 4, 5, 7, 8, and 9. As there are 14 items in the questionnaire, the possible number of missing patterns is $2^{14} = 16384$. However, these data with 360 cases showed only 56 patterns. From the figure it is clear that no single pattern is prominent and the occurrence seems to be completely random.

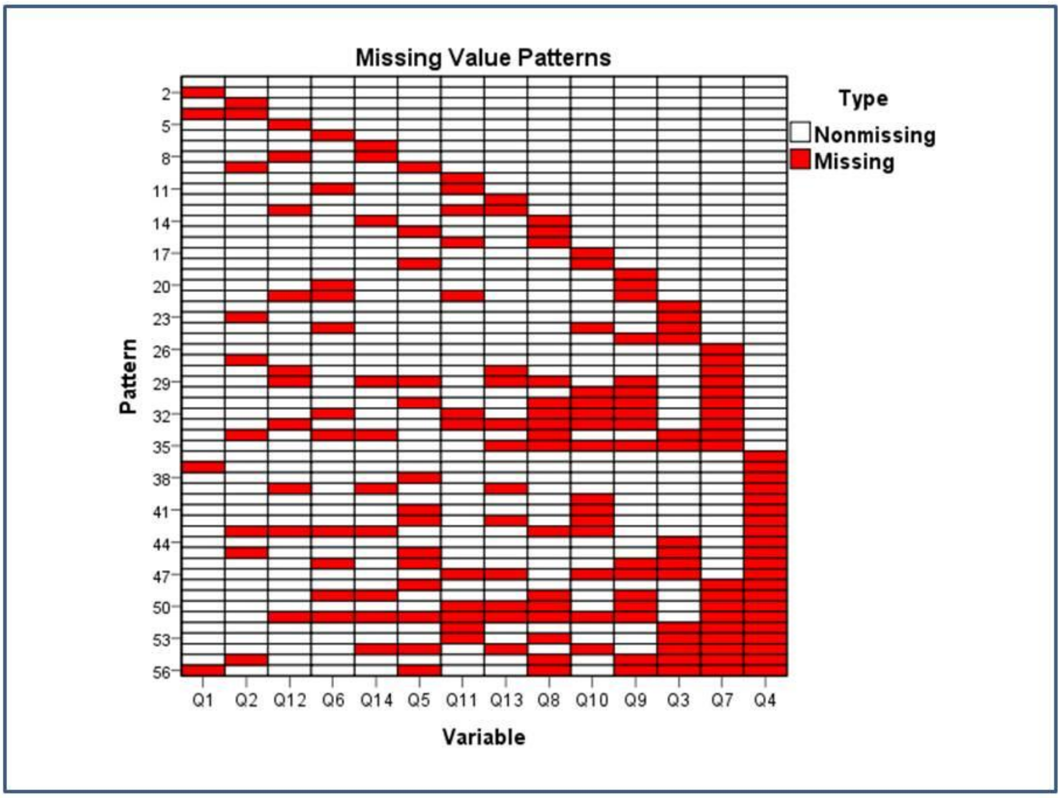

Figure 4.3 Missing data patterns in the data

4.4.2 Comparison of Methods

As only 76.94% of the samples provided complete data to all the 14 items, the missing data in the total score was 23.06% which is considerably a high percentage of missing data. Therefore, various methods to handle the missing data were investigated and these methods were compared. The missing data were filled with item mean, subject mean, interpolation, regression method, using trend in the data, Expectation Maximization (EM) algorithm and Multiple Imputation (MI) methods as described above. The data created with these different methods were analysed for comparison between the methods apart from complete case analysis. Mean, standard deviation and skewness were used to compare these methods. Table 4.3 summarises these measures for different methods of imputation of missing values.

Table 4.3 Estimates using different missing data techniques

Method	Mean	Standard deviation	Median	Range	Skewness
Complete case	22.81	16.10	20.0	61	0.45
Item mean	21.78	15.30	19.0	61	0.61
Subject mean	21.46	15.95	18.0	66	0.60
Interpolation	21.94	15.37	19.0	61	0.59
Regression	21.68	15.72	19.0	62	0.60
Trend	21.78	15.31	19.0	61	0.61
EM Algorithm	21.63	15.75	19.0	62	0.60
Multiple Imputation	21.75	15.67	19.0	65	0.60

Among the eight missing data handling methods tested, the complete case analysis showed highest mean and standard deviation values, suggesting the possible over estimation of the mean and median scores. However, the skewness for this method was the lowest among all methods. As far as the mean score is concerned, all the methods gave similar scores and the imputation by subject mean was the lowest.

4.4.3 Selection of suitable method for handling missing data in OHIP-14 items

As this study has a sufficiently large sample size, the question of loss of power due to missing data does not arise. Moreover, the missing data in OHIP items was shown to follow MCAR pattern and the internal consistency of OHIP-14 is above 0.90. In such situations, it is common that all the above missing data handling methods tend to produce similar results. From Table 4.3, it is obvious that all the methods produce similar mean values for the total OHIP score. However, as this study is mainly to compare different methods and to suggest a suitable method to handle missing data in OHIP-14 items, comparison of various methods both with the above parameters and also the validity of the methods have been discussed.

Currently, the method of Multiple Imputation is commonly used and this method imputes the missing values by adopting a suitable model and the imputed values are again used to estimate the value in the next step. The estimates obtained using multiple imputation have less bias and closer to the original values. As increasing number of imputations reduces the difference between the original and estimated values, in this study a total of 10 imputations were performed using the regression model. Based on the multiple imputed data sets, the mean total score was 21.75 with a standard deviation of 15.67. The median score was 19 and the skewness was 0.60.

The performance of the missing data methods are compared by using the estimates of the parameter and the standard deviation/ standard error (Dong and Peng, 2013). In the present analyses, all the eight methods produced similar estimates for the mean value and also for the standard deviation. Under MAR assumption, methods based on Maximum Likelihood such as MI and EM algorithm provides valid inferences (Dong and Peng, 2013). The advantage of using Multiple Imputation is its generalizability and replicability. The practical difficulty in analysing the imputed data set is the pooling of results for each imputed data set. Though software provides the pooled measures, it is not available for all types of analysis especially when composite scores are created using multiple items as in this case. In such, circumstances, the missing data handling by EM algorithm can be used which is equally good in terms of obtaining estimates closer to the original values.

4.4.4 Missing data in the follow-up study

All the participants (n=360) at baseline were also asked to complete the OHIP-14 questionnaire after 2 months and 4 months following initial data collection. In the first follow-up, a total of 89 (24.72%) participants completed the questionnaire and a total of 75 (20.83%) participants completed the second follow-up questionnaire. 58 (16.11%) of the participants provided data at all three time points.

The mean and standard deviation of total OHIP-14 scores based on all the 14 items at various follow-ups for the demographic variables are summarized in Table 4.4. It is clear from the table that during follow-up periods the mean OHIP scores decreased, suggesting that the Oral Health Related Quality of Life of these patients tended to increase. The mean total score was consistently higher for females than males at all three time points indicating that females experienced poorer OHRQoL when compared to males throughout the study period. Similarly, the mean score is lower for 'A levels' and 'Graduate' patients than the 'GCSE' and 'others' (diploma and certificate courses) at all three time points. This suggests that educated patients experience better OHRQoL than GCSE than those with lower educational attainment.

Table 4.4 Mean and standard deviation of OHIP scores at various follow-up points by demographic details.

Variables	Baseline		Follow-up-1		Follow-up-2	
	N	Mean (sd)	N	Mean (sd)	N	Mean (sd)
Sex						
Male	148	20.07 (15.52)	30	18.40 (17.37)	29	15.28 (16.42)
Female	212	22.71 (15.85)	59	21.61 (16.11)	46	20.52 (16.62)
Ethnicity						
Whites	246	21.74 (15.71)	67	20.01 (16.99)	59	18.64 (17.16)
Black	59	20.80 (15.78)	10	22.40 (17.48)	9	18.89 (19.50)
Mixed	30	22.63 (17.03)	6	16.00 (8.76)	4	13.25 (7.89)
Asian	17	21.24 (15.71)	5	30.40 (16.62)	3	21.33 (4.73)
Others	4	24.50 (14.48)	-	-	-	-
Education						
GCSE	77	24.98 (15.99)	17	24.29 (16.52)	14	23.64 (16.53)
A levels	50	18.84 (13.25)	11	16.91 (9.77)	11	15.73 (15.04)
Degree	161	20.16 (14.71)	40	17.38 (14.86)	32	13.88 (11.96)
Others	39	28.28 (19.55)	14	25.86 (21.10)	10	26.70 (24.12)

4.4.5 Exploring dropout mechanisms

The mean and 95% confidence interval of the total score for dropout and non-dropout groups at various time points are shown in Figure 4.4. From the figure it is clear that the confidence intervals overlap between dropouts and non-dropout groups indicating that the difference between dropouts and non-dropouts with respect to the mean total scores is not significant. Therefore, it may be concluded that the earlier values do not predict the dropouts.

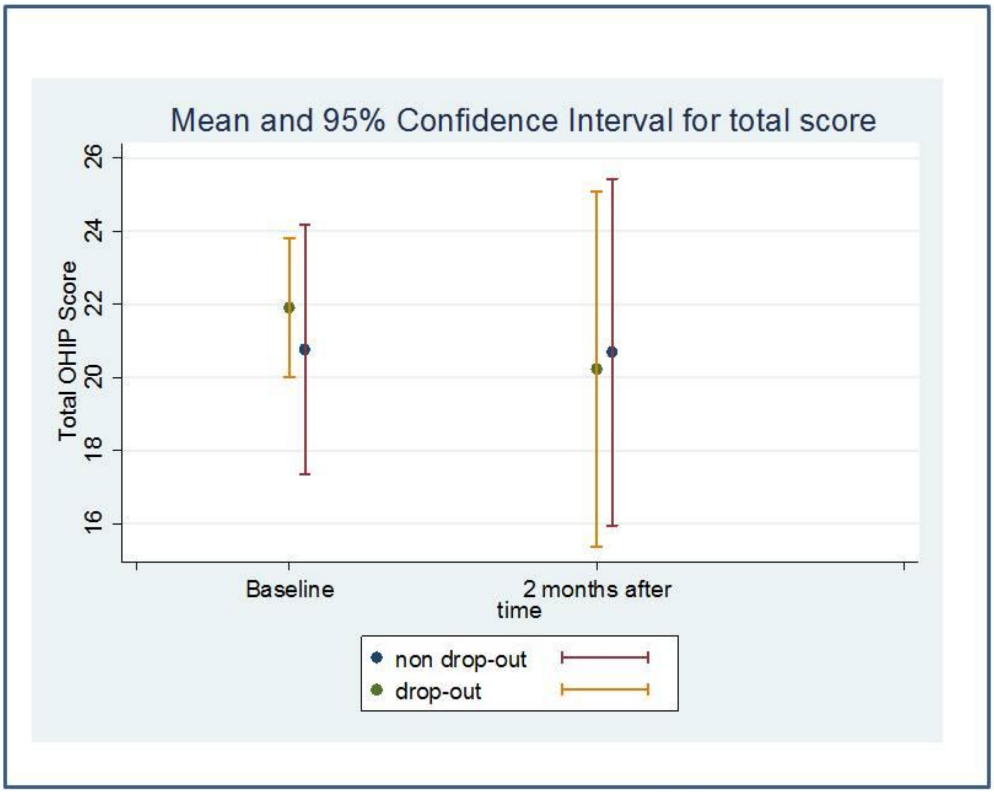

Figure 4.4 Mean and 95% CI for dropouts and non-dropouts.

To determine whether the subject characteristics predicted the dropouts, a logistic regression model was fitted with the demographic variables as the predictors and dropout status as the outcome measure. The results of the logistic regression are summarized in Table 4.5. From the table, it is observed that age, gender and relationship significantly predicted the dropouts. This clearly shows that earlier OHIP values did not predict the dropouts but the subject characteristics did predict the dropouts, indicating that there is less chance of dropouts following a Missing Completely At Random (MCAR) pattern. Hence, there is a high chance of the missing data following the pattern of

Missing At Random (MAR) in which earlier values or the subject characteristics do predict the dropouts. Therefore, it is important to consider an analysis that provides valid results even if the subject characteristics (age, gender and relationships) predicted the dropouts.

Table 4.5 Results of Logistic regression model to find out whether subject characteristics predict the dropouts.

Predictor	P value
Age	<0.01*
Gender	0.04*
Ethnicity	0.18
Education	0.16
Relationship	<0.01*
Profession	0.99
Needs	0.57

*denotes significant predictors of dropouts.

The main analysis was carried out to determine whether there is a significant change in the overall mean OHIP score between three time points (baseline and two follow-ups) and to determine whether any other characteristics predict the mean composite OHIP score. Random Effects Models were fitted by including age, gender and relationship as they predicted the dropouts along with time and other potential predictors (education, profession and treatment needs).

In order to increase the power of this analysis, univariate models were tested for each of the potential predictors along with age, gender and relationship and the variables that significantly predicted the total score at a liberal 10% level were then included in the final model. The results of the univariate models for each predictor are summarised in Table 4.6.

Table 4.6 Results of individual Random Effects Models to find out the significant predictor

Predictor	P value
Ethnicity	0.97
Education	<0.01
Profession	0.10
Treatment Needs	0.05

As Education, profession and treatment needs were the significant predictors at a liberal 10% level in the univariate models, these variables were included in the final model. Hence, the final model included Total OHIP score as the outcome variable and Age, Gender, Relationship, Education, Profession and Needs as the predictor variables. In order to test the moderation effect of treatment needs, the initial model included interaction between time and needs. As the interaction effect was not statistically significant (p=0.25), the final model included only the main effects. The results of the final Random Effects Model used to find the temporal effect of OHIP total scores are given in Table 4.7.

Table 4.7 Results of Random Effects Models

Predictors	**Compared with**	**Effect**	**Z score**	**P value**	**95% Confidence Interval**	
					LCL	UCL
Age		0.09	1.01	0.31	-0.08	0.26
Gender	Male					
Female		0.15	0.07	0.94	-3.87	4.18
Relationship	Single					
In a relationship		0.14	0.06	0.96	-4.94	5.23
Married		-2.34	-0.85	0.40	-7.75	3.07
Separated and other		-7.97	1.84	0.07	-0.39	12.57
Time	Baseline					
Follow-up-1		-0.90	-0.65	0.52	-3.62	1.82
Follow-up-2		-2.41	-1.69	0.09	-5.21	0.39
Education	GCSE					
A level		-7.97	-2.37	0.02*	-14.57	-1.37
Degree and higher		-5.25	-2.00	<0.05*	-10.38	-0.11
Diploma and other		2.26	0.60	0.55	-5.14	9.65
Profession	Unemployed					
Full time		-8.42	-2.33	0.02*	-15.49	-1.34
Part time		-7.14	-1.72	0.09	-15.30	1.02
Student		-11.70	-1.86	0.06	-24.05	0.65
Retired		-17.93	-3.90	<0.01*	-26.94	-8.92
Other		-7.16	-1.51	0.13	-16.45	2.13
Treatment Needs	Restorative					
Orthodontic		-3.51	-1.14	0.25	-9.52	2.50
Operative and surgery		0.05	0.02	0.99	-5.15	5.25
General check-up		-3.44	-0.49	0.62	-17.18	10.30
Multiple needs		10.45	3.24	<0.01*	4.12	16.78

*denotes statistically significant. Negative scores indicate lower score for these groups and hence better OHRQoL.

The results showed that although there was a decrease in total OHIP score between the baseline and follow-up periods, it was not statistically significant. Age, Gender and relationship did not significantly predict the total OHIP score. However, the score differed significantly according to education, profession and treatment needs ($p<0.05$). People with A levels and a degree or higher educational qualifications had significantly ($p=0.02$ and $p=0.04$ respectively) lower scores when compared to people with GCSE only, indicating that they experienced better OHRQoL. Among the different employment categories, 'Full time' and 'Retired' people had significantly ($p=0.02$ and $p<0.01$ respectively) lower scores than the unemployed group. Similarly, people who had multiple treatment needs had significantly ($p<0.01$) higher scores than people with restorative treatments indicating that people with multiple treatment needs experience worse OHRQoL than others.

The predicted mean and 95% confidence interval for the three time points based on the Random Effects Model are shown in Figure 4.5. It is clear from the figure that people with multiple needs had significantly higher total scores than other groups at all three time points. However, within each group, none have recorded significant increase/ decrease in scores between the time points. People who had a general check-up recorded the lowest score at all three time points.

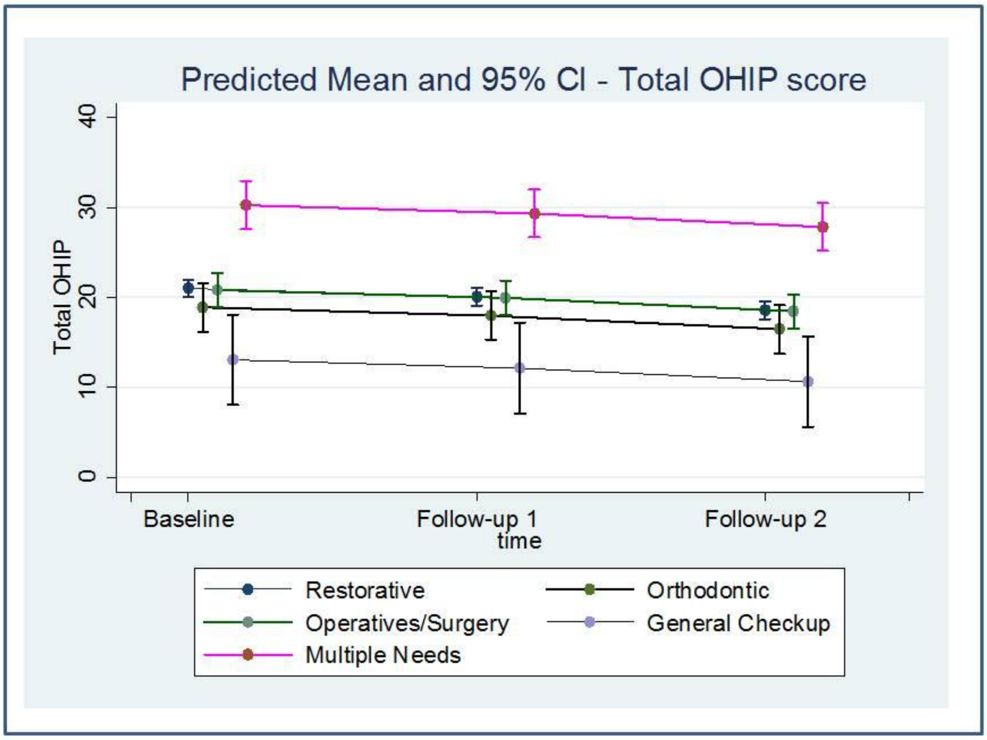

Figure 4.5 Predicted mean and 95% confidence interval at three time points for different treatment groups.

4.5 Discussion

In this chapter, the data have been analysed to address the problems associated with missing data in surveys such as that obtained by OHIP-14. In the dataset collected for this study, item missingness followed a Missing Completely At Random pattern, indicating that the missing data in OHIP items was not due to any specific reasons but occurred at random. This finding facilitates using complete case analysis, whereby missing records are ignored for further analysis.

Although complete case analysis is efficient when the sample size is large, in real life, researchers tend to utilise only the required sample at the beginning of the study using a prior power calculation. This can readily lead to an insufficient sample size at the end of the study after removing the missing records. In such a situation an alternative procedure is required in order to maintain the power of the study. Though some studies have used mean imputation or regression methods (Guarnizo-Herreño et al., 2014, Saub et al., 2005) to handle missing data, the use of these methods mainly depends on the sample size and the power of the study. The imputation of missing data by various methods in this study showed that complete case analysis had slightly higher mean values when compared to other methods. Due to the advancement in computing power, the use of multiple imputation in handling missing data has been widely used by researchers. The use of complete case analysis and mean substitution in handling missing data is common in the literature not only in Oral Health research but also in other areas. However, use of these methods have been repeatedly questioned by others (Rubin, 1987, Schafer, 1997).

When the missing data are assumed to follow MAR pattern and the data set contains large number of missing data, the imputation methods based on Maximum Likelihood are suggested (Dong and Peng, 2013) and particularly the use of auxiliary variables in multiple imputation methods can be efficient as they reduce bias and improve precision of estimates (Collins et al., 2001, Graham, 2009, Hardt et al., 2012). Though both MI and EM algorithm are Maximum Likelihood based methods, the MI method involves severe calculation and requires large memory usage both of which adds considerably to the amount of time required for this kind of analysis. The majority of the standard statistical software packages provide imputation of data using both of these methods but advanced analyses are limited for data created using MI. Therefore, the EM method may be preferred when handling missing data in this context. Data imputed with the EM algorithm may be treated as the original data with missing data filled and hence all further analyses can be carried out with the software. In the present study, though both methods produced similar results, the use of EM algorithm to handle missing data is suggested as this method is unbiased and efficient under MAR condition, and also simple and easy to implement (Dong and Peng, 2013).

The overall mean OHIP score indicates that the dental problems do affect the OHRQoL of both men and women equally as described in the previous chapter. Although many things like food/ eating habits and daily activities may differ between different ethnic groups, the OHIP score did not differ between different ethnic groups. But the total OHIP score differed significantly between different occupations. It is interesting to note that 'Retired' people experienced better OHRQoL when compared to unemployed. A study has shown that unemployed people have a flexible daily routine which hinders them from maintaining regular and beneficial Oral Health Related Behaviour activities and was concluded that unemployed people could be a risk group for poor oral-health related behaviours (Al Sudani et al., 2016). As these data are highly heterogeneous in terms of dental conditions and the samples are all self-selected Dental School patients, it is difficult to demonstrate the external validity. The high rate of dropouts during the follow-up may be due to the dental conditions the patients are suffering from at the time of recruitment in this study.

Based on the current analyses, it is observed that the missing data in OHIP-14 items follow an MCAR pattern and in the follow-up, they followed a MAR pattern. All the eight missing data handling methods considered in this study provided similar results. This study had a reasonably large sample size, which may be the reason for similar results (estimates) obtained with all the methods. However, since the majority of the research studies involve smaller samples, missing data in the data set will have greater impact on the analysis results (estimates). In such situation, methods based on ML will provide better estimates. As the pattern of missing data is more important, under MAR condition, the use of either Multiple Imputation method or EM algorithm is suggested to impute the missing data.

5 Floor and Ceiling effects of the OHIP-14

5.1 Chapter introduction

This chapter aims to identify the floor and ceiling effects and incorporate these findings in studying the relationship between various factors and OHRQoL scores. In assessing health status scores, one common problem researchers encounter in the data is the presence of floor and ceiling effects. As the instruments used for health status assessment often use Likert-type scales ranging from 1 to 10 for some questions, it is possible to have lowest and highest values for some of the questions. All the patients who gave the lowest scores may not have the same OHRQoL and similarly with the patients who scored highest. Most of the instruments used for oral health assessment use scales ranging from 1 to 5. As the range becomes smaller, the probability of having ceiling and floor effects is higher (Hyland, 2003). These differences will therefore have the effect of decreasing the discriminatory usefulness of such an assessment. Hence, it is essential to create awareness amongst researchers about using these measures because of the problems in interpreting these results when measuring the health status.

The floor effect is where many individuals score the lowest value for a particular item. As an example, all individuals who answer 0 (lowest value, which means strongly disagree) for the question “Tooth pain does not affect eating” may not all have the same OHRQoL due to tooth pain. As there are no further scoring options to assess the poor OHRQoL, the individuals are forced to give the same answer. This will significantly affect the overall measure of Oral Health Related

Quality of Life measure and in particular, may make it more difficult to measure a decline in oral health score in other words difficult to measure an improvement in oral health status from a reasonably healthy baseline. The ceiling effect is very similar to the floor effect and occurs when many individuals give the highest score for a particular item. This limits the researcher's ability to measure the impact of treatment interventions on change in health status or measuring the change scores. Consequently, this chapter investigates the presence of floor and ceiling effects in the study dataset and their effects.

5.2 Methods

In addition to the descriptive statistical summaries, multiple linear regression was used to assess the relationship between OHIP total score and demographic variables. In a conventional way, the ordinary least square (OLS), without considering floor or ceiling effects, was applied. In addition, the Tobit regression was used by adjusting any potential floor and ceiling effects present in the data. The results of these two regression models were compared using the difference in the effect of predictors and the standard errors.

The initial development of the Tobit model was based on econometrics data by Tobin (Tobin, 1958). According to Tobin, the linear regression model with censored data can be represented as:

$$y^* = a + b^*x + e$$

where,

- y^* is the latent variable to be measured based on the observed data (x)
- 'a' and 'b' are the constants to be estimated and
- 'e' is the normally distributed error term.

The observed true score is y and,

$$y = y^* \quad \text{if } y^* < c$$

$$y = c \quad \text{if } y^* \geq c;$$

The observed score y is measured with right-censoring and y cannot exceed the censoring point c.

As OHIP-14 uses Likert scales, the actual score is expected to be biased due to censoring. The possibility of both censoring at the lower and the upper end yielding floor and ceiling effects respectively were considered. Both the seven dimensions of the OHIP-14 score, as well as the total score of all the 14 items – the composite score, were considered.

It is usual to assume the lowest score to be the left-censoring point. However, in the case of composite scores like OHIP-14 score, there are two options to determine this:

- to take the total of the lowest scale scores of all the items (in this case the lowest scale score is 0 for all the 14 items and hence the total is 0) which is an extreme point;
- to consider the sum of all the lower end scores of all the items (if the scale range is 0 to 5, maybe 0 and 1 are considered as the lower ends).

As there is no clear definition of lower end scores, there are many possible options according to the range of scales. The method suggested by Ven den Oord and Van den Ark (Oord and Ark, 1997) to decide about the floor and ceiling points in the case of composite data was adopted in this research. The 'average item difficulty' representing the proportion of subjects whose rating is consistently biased towards the lower end of the scale was established. Similarly, for the ceiling point the 'average item easiness', representing the proportion of subjects whose rating is consistently biased towards the upper end is calculated. Since the total OHIP score is discrete, linear interpolation was used to compute the censoring point. According to this method, suppose that the x% of subjects consistently fall into the lower end of the score, and the cumulative distribution of scale scores indicates that 30% of the subjects scored 9 or below and 50% of the subjects scored 10 or below where x lies between 30% and 40%, then the censoring point is calculated as,

$$\text{Censoring Point} = 9 + \frac{(X - 40)}{(50 - 30)}$$

Tobit regression is fitted by adjusting the calculated floor and ceiling points for the composite data.

5.3 Floor and Ceiling Effects

All the items of the OHIP-14 were scaled from 0 to 5, with higher scores indicating poorer OHRQoL. The missing data in the OHIP-14 items were filled using the EM algorithm. The number of patients who recorded the lowest score (floor) and the highest score (ceiling) for each of the 14 items are shown in Figure 5.1.

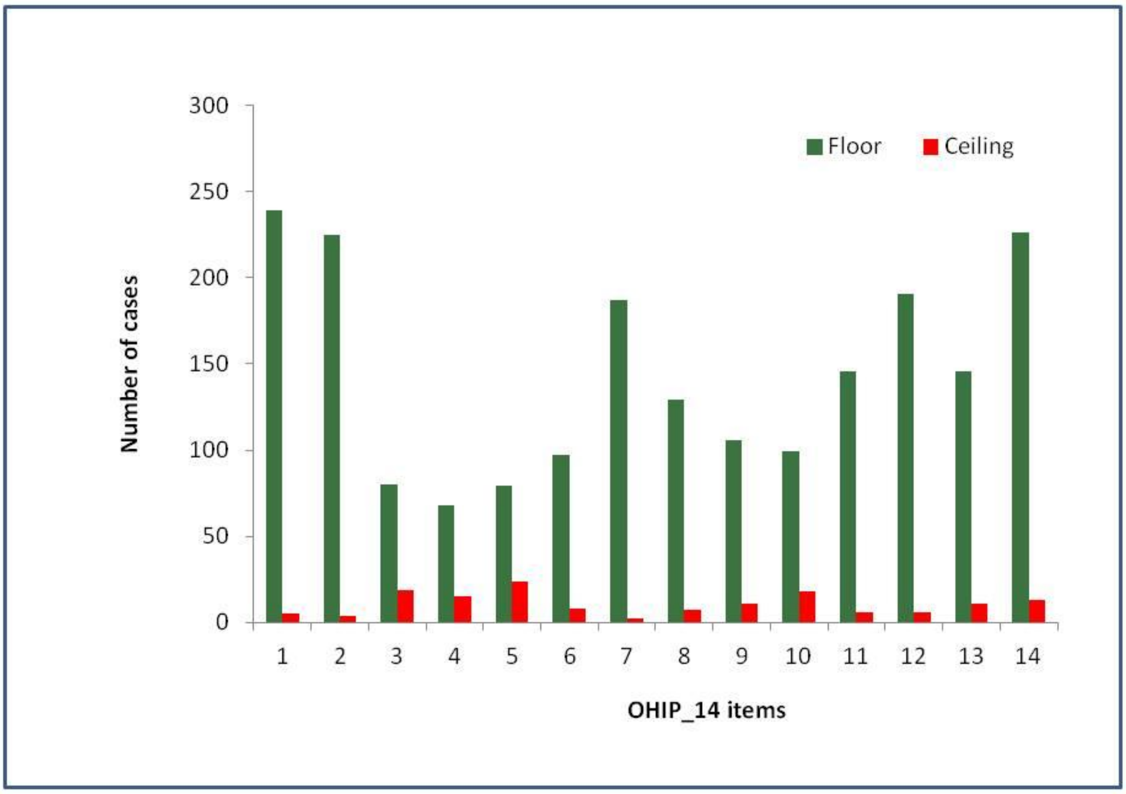

Figure 5.1 Total number of samples who scored lowest (0) (Floor) and highest (5) (ceiling) Scores in each of OHIP-14 items.

It is clear from the figure that the floor score is seen considerably more often than the ceiling score. The amount of left censoring (floor) ranged between 18.90% and 66.40% for the 14 items, while the amount of right censoring (ceiling) ranged between 0.60% and 6.70%. The 'average item difficulty' (percentage of subjects with the lowest item score) was 0.4 indicating that on average 40% of the patients consistently gave the lower end of the score. On the right censoring (ceiling), 'the average item easiness' was 0.03, meaning that 3% of the patients tended to give higher side of the score.

Ceiling point at lower side

In the cumulative distribution of scale scores, 39.70% of the samples scored 14 or lower and 42.80% scored 15 or lower. Hence the censoring point at the lower side is,

$$= 14 + \frac{(40 - 39.70)}{(42.80 - 39.70)} = 14.09$$

Ceiling point at upper side

In the cumulative distribution of scale scores, 2.20% of the samples scores 56 or over and 3.60% scored 56 or over. Hence the censoring point at the upper side is,

$$= 55 + \frac{(3 - 2.20)}{(3.6 - 2.20)} = 55.57$$

Therefore, based on the method suggested by Van den Oord and Van den Ark (Oord and Ark, 1997), the censoring points at lower side (floor) and higher side (ceiling) for the composite score of OHIP-14 were 14.09 and 55.57 respectively.

Table 5.1 Results of OLS and Tobit model fitting.

Predictors	Ordinary Least Squares (OLS) Regression		Tobit Regression	
	Effect (95% CI)	p value	Effect (95% CI)	p value
TOTAL OHIP-14				
Age	-0.02 (-0.15, 0.12)	0.83	-0.02 (-0.20, 0.16)	0.83
Gender	1.90 (-1.48, 5.27)	0.27	2.79 (-1.76, 7.33)	0.23
Profession:		0.06		0.06
Full time	-5.70 (-11.09, -0.32)	0.27	-6.20 (-13.27, 0.86)	0.09
Part time	- 5.10 (-11.64, 1.50)	0.13	-6.00 (-14.61. 2.71)	0.18
Student	-10.90 (-20.42, -1.27)	0.03*	-12.90 (-26.10, 0.31)	0.06
Retired	-10.20 (-17.37, -3.01)	0.01*	-14.70 (-24.6, -4.9)	<0.01*
Other	-6.70 (-13.96, 0.54)	0.07	-7.00 (-16.60, 2.64)	0.16
FUNCTIONAL LIMITATION				
Age	0.03 (0.01, 0.04)	0.01*	0.05 (0.01, 0.08)	0.02*
Gender	-0.08 (-0.56, 0.40)	0.74	-0.84 (-1.02, 0.85)	0.86
Profession:		0.43		0.57
Full time	-0.40 (-1.17, 0.37)	0.31	-0.27 (-1.76, 1.22)	0.73
Part time	-0.84 (-1.78, 0.10)	0.08	-1.49 (-3.38, 0.40)	0.12
Student	-0.26 (-1.62, 1.11)	0.71	-0.11 (-2.83, 2.62)	0.94
Retired	-0.85 (-1.87, 0.17)	0.10	-0.98 (-2.93, 0.97)	0.33
Other	-0.77 (-1.80, 0.26)	0.14	-0.96 (-3.00, 1.09)	0.36
PHYSICAL PAIN				
Age	0.01 (-0.01, 0.03)	0.30	0.01 (-0.01, 0.04)	0.30
Gender	0.22 (-0.35, 0.79)	0.45	0.21 (-0.42, 0.85)	0.51
Profession:		0.35		0.27
Full time	-0.23 (-1.15, 0.68)	0.62	-0.32 (-1.33, 0.69)	0.53
Part time	-0.02 (-1.14, 1.10)	0.97	-0.00 (-1.22, 1.23)	0.99
Student	-0.61 (-2.23, 1.02)	0.46	-0.82 (-2.63, 0.99)	0.37
Retired	-1.23 (-2.45, -0.01)	0.04*	-1.41 (-2.76, -0.08)	0.04*
Other	-0.39 (1.63, 0.84)	0.53	-0.46 (-1.81, 0.90)	0.51
PSYCHOLOGICAL DISCOMFORT				
Age	-0.01 (-0.04, 0.01)	0.39	-0.01 (-0.04, 0.01)	0.30
Gender	0.64 (0.02, 1.26)	0.04*	0.79 (0.09, 1.49)	0.03*

Predictors	Ordinary Least Squares (OLS) Regression		Tobit Regression	
	Effect (95% CI)	p value	Effect (95% CI)	p value
Profession:		0.02*		0.01*
Full time	-0.96 (-1.95, 0.03)	0.06	-1.09 (-2.20, 0.01)	0.05
Part time	-0.72 (-1.93, 0.49)	0.24	-0.90 (-2.25, 0.45)	0.19
Student	-2.31 (-4.07, -0.55)	0.01*	-2.79 (-4.83, -0.75)	0.01*
Retired	-1.89 (-3.21, -0.57)	0.01*	-2.25 (-3.74, -0.75)	<0.01*
Other	-1.56 (-2.89, -0.23)	0.02	-1.72 (-3.22, -0.22)	0.03*
PHYSICAL DISBILITY				
Age	0.01 (-0.02, 0.03)	0.66	0.01 (-0.02, 0.04)	0.63
Gender	0.36 (-0.22, 0.94)	0.22	0.42 (-0.31, 1.15)	0.26
Profession:		0.44		0.43
Full time	-0.69 (-1.62, 0.24)	0.14	-0.80 (-1.96, 0.36)	0.18
Part time	-0.55 (-1.69, 0.58)	0.34	-0.56 (-1.97, 0.85)	0.44
Student	-1.09 (-2.74, 0.56)	0.19	-1.23 (-3.32, 0.85)	0.25
Retired	-1.23 (-2.46, 0.01)	0.05	-1.57 (-3.13, -0.02)	0.04*
Other	-0.48 (-1.73, 0.77)	0.45	-0.53 (-2.10, 1.03)	0.50
PSYCHOLOGICAL DISABILITY				
Age	-0.02 (-0.04, 0.01)	0.15	-0.02 (-0.05, 0.004)	0.09
Gender	0.36 (-0.24, 0.96)	0.24	0.40 (-0.30, 1.11)	0.26
Profession:		0.02*		0.02*
Full time	-1.39 (-2.34, -0.44)	<0.01*	-1.57 (-2.68, -0.46)	0.01*
Part time	-1.37 (-2.54, -0.21)	0.02*	-1.60 (-2.96, -0.25)	0.02*
Student	-2.35 (-4.04, -0.65)	0.01*	-1.76 (-4.79, -0.73)	0.01*
Retired	-1.98 (-3.25, -0.71)	<0.01*	-2.30 (-3.81, -0.80)	<0.01*
Other	-1.51 (-2.79, -0.22)	0.02*	-1.66 (-3.17, -0.16)	0.03*
SOCIAL DISABILITY				
Age	-0.02 (-0.04, 0.01)	0.12	-0.03 (-0.06, 0.00)	0.08
Gender	0.43 (-0.14, 1.01)	0.14	0.60 (-0.17, 1.37)	0.12
Profession:		0.03*		0.03*
Full time	-1.12 (-2.05, -0.20)	0.02*	-1.29 (-2.49, -0.08)	0.04*
Part time	-0.89 (-2.02, 0.24)	0.12	-1.00 (-2.47, 0.46)	0.18

Predictors	Ordinary Least Squares (OLS) Regression		Tobit Regression	
	Effect (95% CI)	**p value**	**Effect (95% CI)**	**p value**
Student	-2.32 (-3.96, -0.67)	0.01*	-3.05 (-5.28, -0.82)	0.01*
Retired	-1.75 (-2.98, -0.52)	0.01*	-2.29 (-3.94, -0.64)	0.01*
Other	-1.03 (-2.28, 0.21)	0.10	-1.20 (-2.82, 0.43)	0.15
HANDICAP				
Age	-0.01 (-0.03, 0.01)	0.36	-0.02 (-0.06, 0.02)	0.27
Gender	-0.04 (-0.64, 0.55)	0.89	-0.01 (-0.94, 0.92)	0.98
Profession:		0.19		0.20
Full time	-0.90 (-1.85, 0.04)	0.06	-1.24 (-2.68, 0.20)	0.09
Part time	-0.67 (1.83, 0.48)	0.25	-0.97 (-2.73, 0.80)	0.28
Student	-1.92 (-3.60, -0.23)	0.03*	-3.12 (-5.89, -0.35)	0.03*
Retired	-1.26 (-2.52, 0.00)	0.05	-1.92 (-3.90, 0.06)	0.06
Other	-0.97 (-2.24, 0.30)	0.14	-1.35 (-3.32, 0.62)	0.18

Note: Comparison group for Gender is Male and Profession is Unemployed.

*denotes statistically significant.

The results of the ordinary least squares (OLS) and Tobit regressions (with a floor value of 14.09 and a ceiling value of 55.57) for the total OHIP-14 and for each of the seven dimensions are given in Table 5.1 .

In the analysis of the total OHIP-14 (see Table 5.1), under both the OLS and Tobit models, age (p=0.27 and p=0.23) and gender (p=0.83 for both) were not significant predictors of OHIP total score. The multiple linear OLS regression analysis showed that ‘students’ and ‘retired’ subjects had significantly lower scores (e.g. better oral health related quality of life) in relation to the unemployed group while, in contrast, under the Tobit model, only 'retired' was found to be significantly different from 'unemployed'. The higher magnitude of the regression coefficients obtained under the Tobit model indicates that the estimated effect of predictor variables on the total OHIP score is higher under the Tobit model than the OLS regression model.

In the analysis by dimensions (Table 5.1) in both the OLS and Tobit models, age was found to be a significant predictor of the OHIP-14 dimension Functional Limitation (p=0.01) and gender was found to be significantly associated with the OHIP-14 dimension Psychological Discomfort (p=0.04). Under both models, 'profession' was not significantly associated with the OHIP-14 dimensions Physical Pain and Physical Disability, but it was significantly associated with all the other dimensions. The effect of profession on these other dimensions was also found for the total OHIP-14: 'Retired' was significantly different from 'unemployed'. In all the cases negative coefficients were obtained, showing 'unemployed' to have the worst outlook. As observed, the results of the significance tests for the regression coefficients were very consistent between the two models, OLS and Tobit. However, the estimate of the regression coefficients is much larger under Tobit regression, confirming that the OLS, without considering floor and ceiling effects, under-estimates the effects.

5.4 Model comparisons

Both the regression and Tobit models were compared using log likelihood and model fit statistics. These statistics for the model with total OHIP score and for the dimension scores are given in Table 5.2.

Table 5.2 Model fit statistics for OLS and Tobit models.

Outcome	**Model**	**P value**	**Log likelihood**
Total OHIP Score	OLS Tobit	0.06 0.05	-1448.16 -99.03
Functional Domain	OLS Tobit	0.12 0.17	-767.73 -485.15
Physical pain	OLS Tobit	0.46 0.40	-829.85 -774.38
Psychological discomfort	OLS Tobit	<0.01 <0.01	-856.85 -769.10
Physical disability	OLS Tobit	0.45 0.48	-834.39 -703.00
Psychological disability	OLS Tobit	<0.01 <0.01	-844.33 -741.88
Social disability	OLS Tobit	<0.01 <0.01	-833.09 -686.13
Handicap	OLS Tobit	0.24 0.20	-841.33 -619.21

The p values for the total score and the dimensions Functional limitation, Physical pain, Physical disability and Handicap were greater than 0.05 indicating that both the OLS and Tobit model fit the data. However, for the dimensions Psychological discomfort, Psychological disability and Social disability the p values were less than 0.05 indicating that the model does not fit the data. The log likelihood values for the Tobit model for all the dimensions and for the total score were less than the likelihood values for the OLS model indicating that the Tobit model fits better than the OLS model.

5.5 Discussion

The presence of floor and ceiling effects in Likert scales has long been established (Oord and Ark, 1997). The presence of the floor effect in the OHIP-14 was larger than the ceiling effect, indicating a tendency towards a good OHRQoL for this specific patient population.

Many studies have analysed the relationship between OHIP-14 scores and demographic variables although none seem to have taken floor and ceiling effects into account. Liu and co-workers (Liu et al., 2012), using univariate analysis, showed that neither age nor sex influenced the total OHIP-14 score. Similarly, a weak association between OHIP-14 mean score and age was established by Slade and Sanders (Slade and Sanders, 2011). They found this association to be weak for people with no clinical conditions but, in contrast, a strong three-fold inverse association between age and mean OHIP score was found for persons with two or more clinical conditions. Similarly, no significant association between OHIP scores and demographic characteristics was reported using Fisher's exact test (Macedo and Queluz, 2011). These findings, with and without adjustment for floor and ceiling effects, were in agreement with these results, in the sense that neither age nor sex are significant predictors of the Total OHIP-14 score.

The influence of floor and ceiling effects on the total OHIP composite score is demonstrated in this study by the difference observed between the regression coefficients of both models (OLS and Tobit). The difference in coefficients was maximum for the *retired* group, whereas, in other predictors it was marginal. Therefore, any model fitted without adjusting for floor and ceiling effects when a considerable proportion of cases suffer from these would be likely to underestimate (in terms of magnitude) the effects. Similar results were reported using simulated data set (McBee, 2010). Using empirical and simulated data, it has been shown that the ceiling effect in longitudinal data lead to biased parameter estimation (Wang et al., 2008). The results presented here confirm that ignorance of floor and ceiling effects may lead to misleading results.

In this study, the multiple OLS regression analysis (ignoring floor and ceiling effects) found *full time*, *student* and *retired* persons significantly associated with the Total OHIP-14 score. However, the Tobit model (a model that takes floor and ceiling effects into account) showed only *retired* persons significantly associated. The Tobit model established a significant reduction in OHIP scores for *retired* people indicating better OHRQoL when compared to the *unemployed* category. Compared to 30-49 year olds, less frequent adverse impacts of oral conditions have been reported among those over 70 in the UK and Australian populations (Steele et al., 2004). The findings in this study are in line with others which suggest better subjective oral health in old age (Dahl et al., 2011, Sanders et al., 2009). Similarly, a study carried out in Australian population found no evidence of poorer subjective oral health in older age (Slade and Sanders, 2011).

An alternative approach for analysis of OHIP-14 data in order to try to minimise the influence of floor and ceiling effects, is the categorisation of responses as either 'no impact' and 'impact' groups instead of dealing with the raw scores. Categorising self-perceived OHRQoL measured using OHIP-14 into two groups using median splits (0 and 1 as one category and 2,3 and 4, as another category) has been suggested previously (Locker, 2003). OHIP-14 responses were also analysed by taking different cut off points (0 as one group and all other responses as another groups) (De Oliveira and Sheiham, 2004). The cut off based on the total score has also been reported in the literature (Willumsen et al., 2010). However, the loss of information that occurs due to dichotomisation of OHIP-14 scores and the absence of an established cut off point for the definition of lower OHIP score in the literature have been highlighted previously (Ikebe et al., 2007, Krisdapong and Sheiham, 2014, Niesten et al., 2016, Tsakos et al., 2012). The presence of serious negative consequences in the dichotomisation of data which include loss of information about individual differences has been explored in terms of a statistical perspective and the use of regression and correlation methods without dichotomisation has been suggested (MacCallum et al., 2002). Therefore, the use of regression models in finding the significant predictors of OHIP-14 without categorisation, avoid the loss of information and adjusting for floor and ceiling effects in this model strengthens the findings.

In conclusion, the results presented in this study demonstrated that the data to assess Oral Health Related Quality of Life using OHIP-14 are affected by floor and ceiling effects. Although the impact of floor and ceiling values was not striking, it was verified that the Tobit model fits the OHIP data better than the OLS model. Hence the Tobit model for OHIP data improves the model predictions and fits the data better than the OLS model. Even after adjusting for the floor and ceiling effects, gender and age did not predict the OHIP score. The findings of this study suggest that any future studies on OHRQoL using OHIP-14 should take these effects into account when exploring the relationship between OHIP-14 composite scores and demographic and clinical variables.

6 Factorial structure of the OHIP-14 instrument

6.1 Chapter introduction

The objective of comparing and confirming the number of factors in OHIP-14 was addressed using one, three, six and seven factor models on the study data. Grouping of items and establishing the factorial structure has been a routine process during instrument development for measuring a latent variable. This not only helps researchers in the process of item reduction while measuring the latent variable but also helps to understand the instrument and the measuring process. Item reduction using factor analysis has been used in many fields such as behavioural and social sciences, medicine, economics, and geography as a result of the technological advancements of computers but the advantage of using this technique is limited (Yong and Pearce, 2013). The original developer of OHIP-14 (Slade, 1997) has shown that the 14 items in OHIP-14 form 7 factors and each factor contains 2 items. Though they identified 7 factors in OHIP-14, only two factors are heavily loaded in the factor analysis. They concluded that the items are heavily correlated and have suggested the need for further exploration of the factorial structure. Furthermore, studies carried out in different countries based on OHIP-14, have come up with different factorial structures. In this respect, most of the studies have used factor analysis to test this, which is based on the correlation between the items included in the analysis.

The correlation between the factors is evident from the original study by (Slade, 1997). Hence, the construct validity of different models can be tested by carrying out Confirmatory Factor Analysis (CFA) using Structural Equation Modelling (SEM). CFA is preferred to Exploratory Factor Analysis as this is theory based and the number of underlying factors is hypothesised *a priori* (Kamphaus and Frick, 2005, Keith, 1990). Moreover, this technique helps to understand the causal relationships among the user defined variables and hence better analytical features. In this chapter, the multi-dimensional construct of OHIP-14 which was derived from the theoretical frame work of Locker's conceptual model was tested using SEM based on one, three, six and seven factor models, which have been proposed in the literature while measuring OHRQoL (Baker et al., 2008a, Montero et al., 2010, Santos et al., 2013, Slade, 1997).

6.2 Methods

Though OHIP-14 has been extensively reviewed in different populations (Slade and Sanders, 2011), the construct and discriminant validity of various models were tested in this study. The internal consistency of the items was checked using Cronbach's alpha. The total item correlation was also calculated to ascertain the internal consistency of items in the OHIP-14 questionnaire. The correlation between the items was checked using inter item correlation for any item redundancy.

The models considered in this analysis were:

1. One factor model where all the 14 items load onto one factor (Santos et al., 2013).

2. Three factor model defined by (Montero et al., 2010) with the three dimensions being Psychosocial Impact, which contained 8 items; Pain and Discomfort containing four items; and Functional Limitation with two items.
3. The six factor model derived by (Baker et al., 2008a) with the six factors being Functional Limitation, Pain, Physical Disability, Social Disability, Handicap - each with two items and Psychosocial Impact with four items
4. The seven factor model derived by (Slade, 1997) with the seven factors being Functional Limitation, Pain, Physical Disability, Social Disability, Handicap, Psychological Disability and Discomfort - each with two items.

One Factor Model

The one factor model grouped all the 14 items into a single factor. This model is testing against a multi-dimensional construct of OHIP-14 based on Locker's theoretical framework. The model with the estimated parameters is shown in Figure 6.1.

The Three Factor model

Though many researchers have proposed that there are seven factors in OHIP-14, some researchers have questioned the presence of these. While studying the dimensional structure of OHIP-14 in Spanish workers, it was reported that the items contained three factors, which are identified as Functional Limitation, Pain-discomfort and Psychosocial impacts (Montero et al., 2010). According to them the allocation of OHIP-14 items to one of these three factors is as below (Table 6.1)

Table 6.1 Allocation of items to different factors in the 3 Factor Model (Montero et al 2010)

Factor Number	Factor	Item No	Item description
1	Functional Limitation	1 2	Trouble pronouncing words Worsening Taste
2	Pain-discomfort	3 4 7 8	Painful aching Uncomfortable to eat Unsatisfactory diet Interrupt meals
3	Psycho-social Impact	5 6 9 10 11 12 13 14	Feel Conscious Tense Difficult to relax Embarrassed Trouble getting on with others Difficulty doing usual jobs Less Satisfying Unable to Function

The Six Factor Model

The construct validity of the original OHIP with 49 items has been examined (Baker et al., 2008a) by using Confirmatory Factor Analysis and evaluated the fit using various indices available in Structural Equation Modelling. It was reported that the original seven factor model as initially defined (Slade, 1997) did not fit the data and they proposed a six factor model comprising 22 of the original items in OHIP-49. They concluded that the revised 22 item six factor model was a better fit for their data and emphasized the need for further testing of the scale.

The six factors identified by Baker and co-workers (Baker et al., 2008a) contained 22 items and were derived from OHIP-49. As this contained only 8 items of OHIP-14, the remaining 6 items were allocated to one of the six factors accordingly based on the original allotment in the seven factor model identified by Slade (Slade, 1997) and the extra items were excluded. The allocation of OHIP-14 items according to the six factor model suggested by Baker and co-workers is given below (Table 6.2).

Table 6.2 Allocation of items to different factors in the 6 Factor Model

Factor Number	Factor	Item No	Item description
1	Functional Limitation	1 2	Trouble pronouncing words Worsening Taste
2	Pain	3 4	Painful aching Uncomfortable to eat
3	Psychological Impact	5 6 9 10	Feel Conscious Tense Difficult to relax Embarrassed
4	Physical disability	7 8	Unsatisfactory diet Interrupt meals
5	Social Disability	11 12	Trouble getting on with others Difficulty doing usual jobs
6	Handicap	13 14	Less Satisfying Unable to Function

The Seven Factor Model

The seven factor model was derived in such a way that each factor contained two items (Slade and Spencer, 1994). The allocation of items to the seven factors is as below (Table 6.3).

Table 6.3 Allocation of items to different factors in the 7 Factor Model

Factor Number	Factor	Item No	Item description
1	Functional Limitation	1 2	Trouble pronouncing words Worsening Taste
2	Physical Pain	3 4	Painful aching Uncomfortable to eat
3	Psychological Discomfort	5 6	Feel Conscious Tense
4	Physical disability	7 8	Unsatisfactory diet Interrupt meals
5	Psychological Disability	9 10	Difficult to relax Embarrassed
6	Social Disability	11 12	Trouble getting on with others Difficulty doing usual jobs
7	Handicap	13 14	Less Satisfying Unable to Function

CFA was carried out using SEM separately for each of the models and the validity of the models were compared. CFA is the measurement model in SEM which has two stages, namely: measurement model and structural model. The allocation of items to various factors in each model and its factor loading were shown using the diagram drawn in AMOS 24.0. All the models were examined with Maximum Likelihood Estimation (MLE) using the study data. The model fit was evaluated for each model using various statistics namely chi-squared statistics with p value, Root Mean Squared Error of Approximation (RMSEA) with 90% confidence interval, and the comparative fit indices Tucker-Lewis Index (TLI), Comparative Fit Index (CFI) and Normed Fit Index (NFI). The accepted model fit included non-significant chi-square, RMSEA values less than or equal to 0.08, the values of CFI, TLI and NFI equal to 0.90 or above and the chi-square/df ratio less than 3.0.

6.3 Results

All the four different models were tested using Confirmatory Factor Analysis. The fitted one factor, three factor, six factor and seven factor models are shown in shown in Figure 6.1 to Figure 6.4.

The model fit indices for the corresponding models are given in Table 6.6. None of the four models fitted the data well. The one factor, three factor and six factor models did not meet any of the seven model fitting criteria (p value, Chi-square/df, RMSEA, CFI, TLI, NFI and CMIN/df). However, the seven factor model met only one of the seven model fitting criteria. This indicates that the original 7 factor structure of the OHIP-14 proposed by Slade is not supported by this current data. The next step was to examine the factor correlations to assess the discriminant validity. Accordingly, factor correlations that exceed the value of 0.85 in the confirmatory factor analysis are considered as having poor discriminant validity (Brown, 2006). In the three factor model, one correlation that is between Pain-Discomfort and Psychosocial Impact exceeded the value of 0.85 (r=0.87). In the six factor model, four correlations met this criterion: Pain – Physical Disability (r=0.91), Psychosocial Impact – Social Disability (r=0.86), Psychosocial Impact – Handicap (r=0.94) and Social Disability – Handicap (r=0.99). In the seven factor model, nine correlations had a value greater than or equal to 0.85. They were: Physical Pain - Physical Disability (r=0.91), Psychological Disability (0.89); Psychological Discomfort – Psychological Disability (1.10), Handicap (0.91); Physical Disability – Psychological Disability (0.94), Handicap (0.85); Psychosocial Disability – Social Disability (0.96), Handicap (1.04); and Social Disability – Handicap (1.00). These indicate the constructs in these models are not distinct and hence have poor discriminant validity. Some of the correlations above exceeded 1.00 indicating that the model ends with Heywood's case.

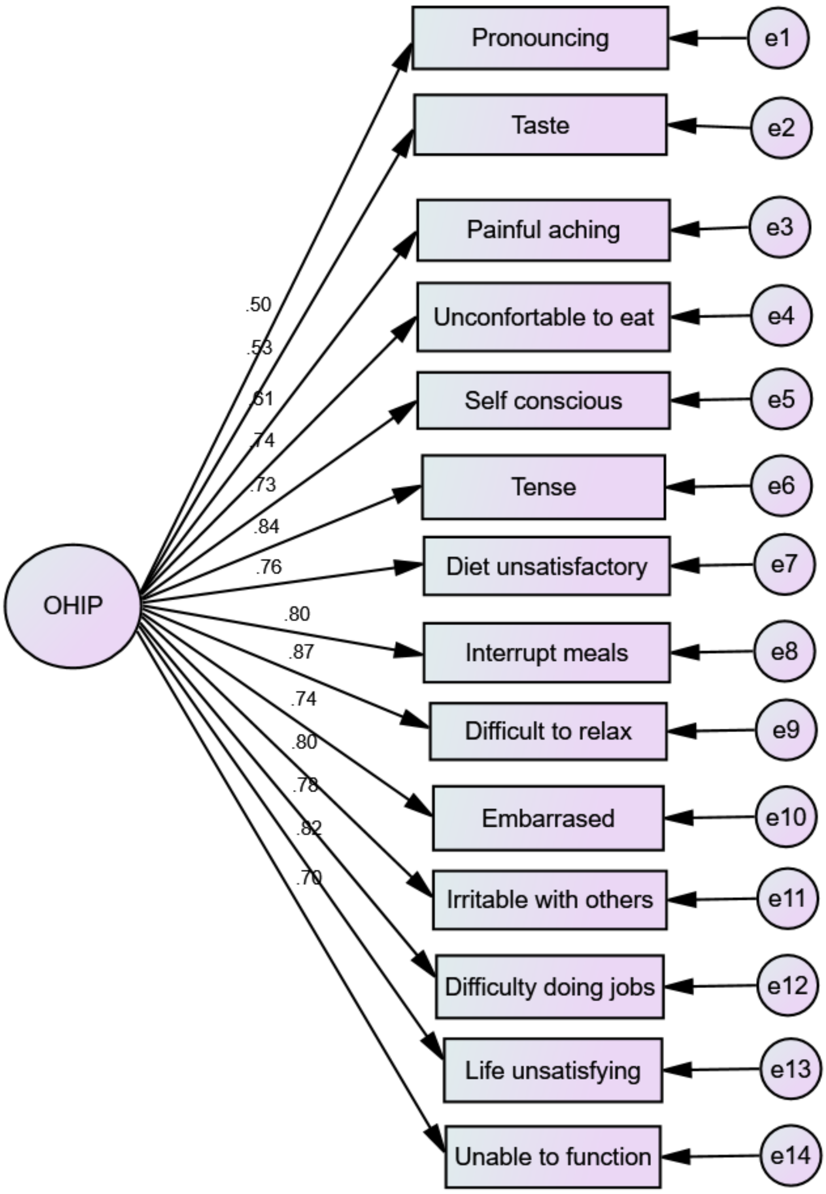

Figure 6.1 Fitted Single factor Structural Equation Model for OHIP-14 items. depicting the relationship between exogenous variables (factors) and endogenous variables (OHIP items) along with covariances.

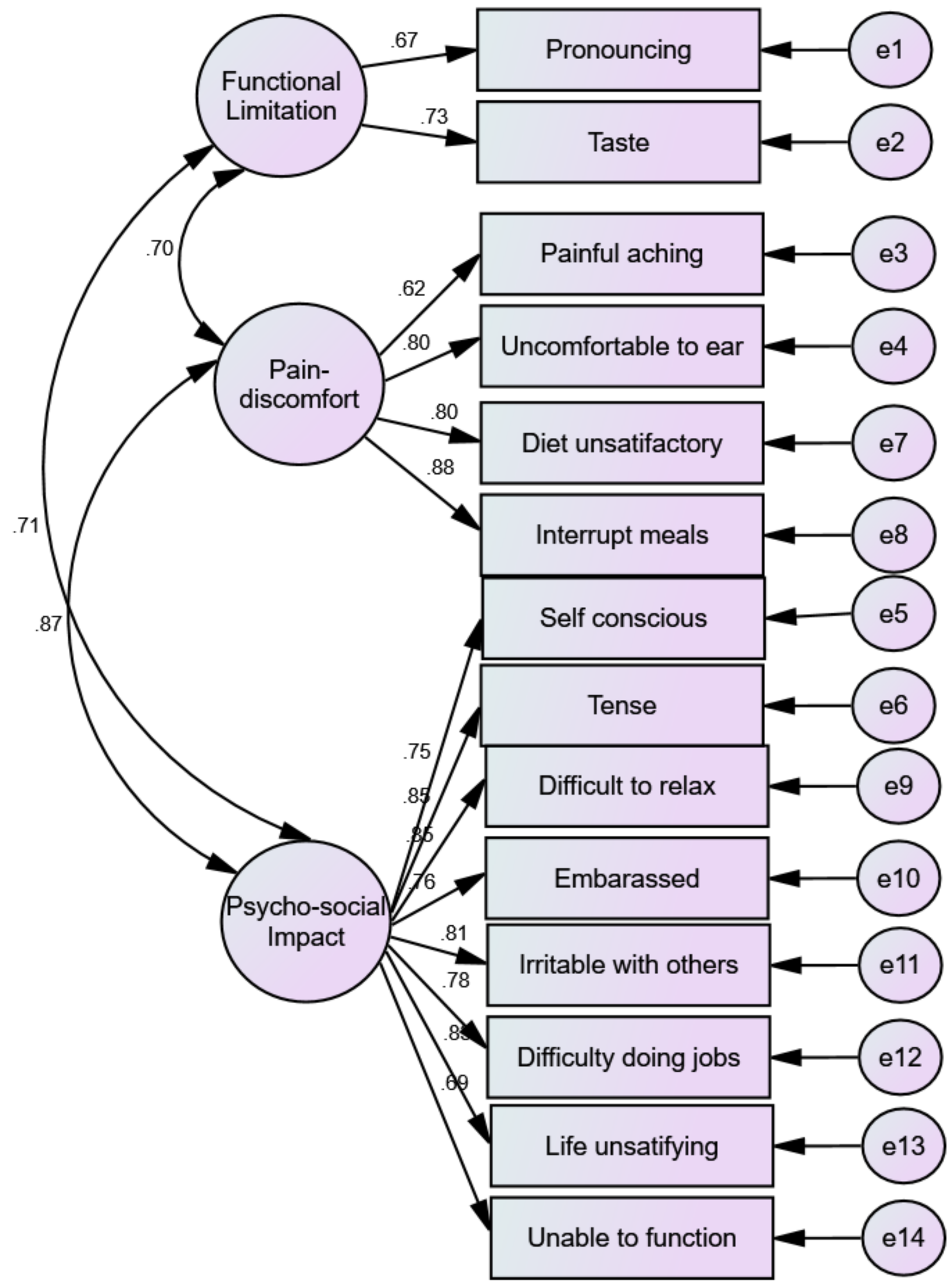

Figure 6.2 Three factor model with estimates

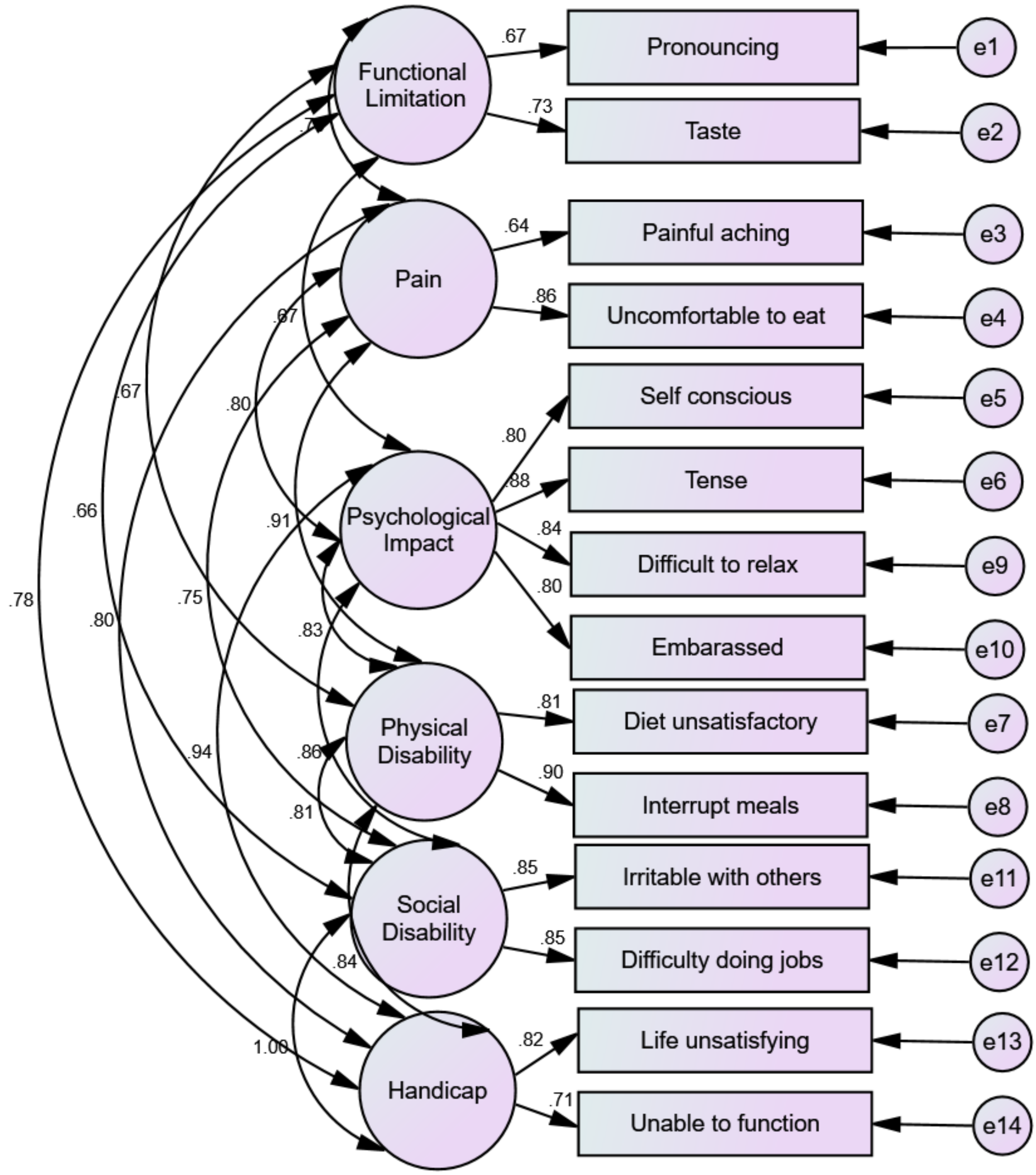

Figure 6.3 Six factor model with estimates

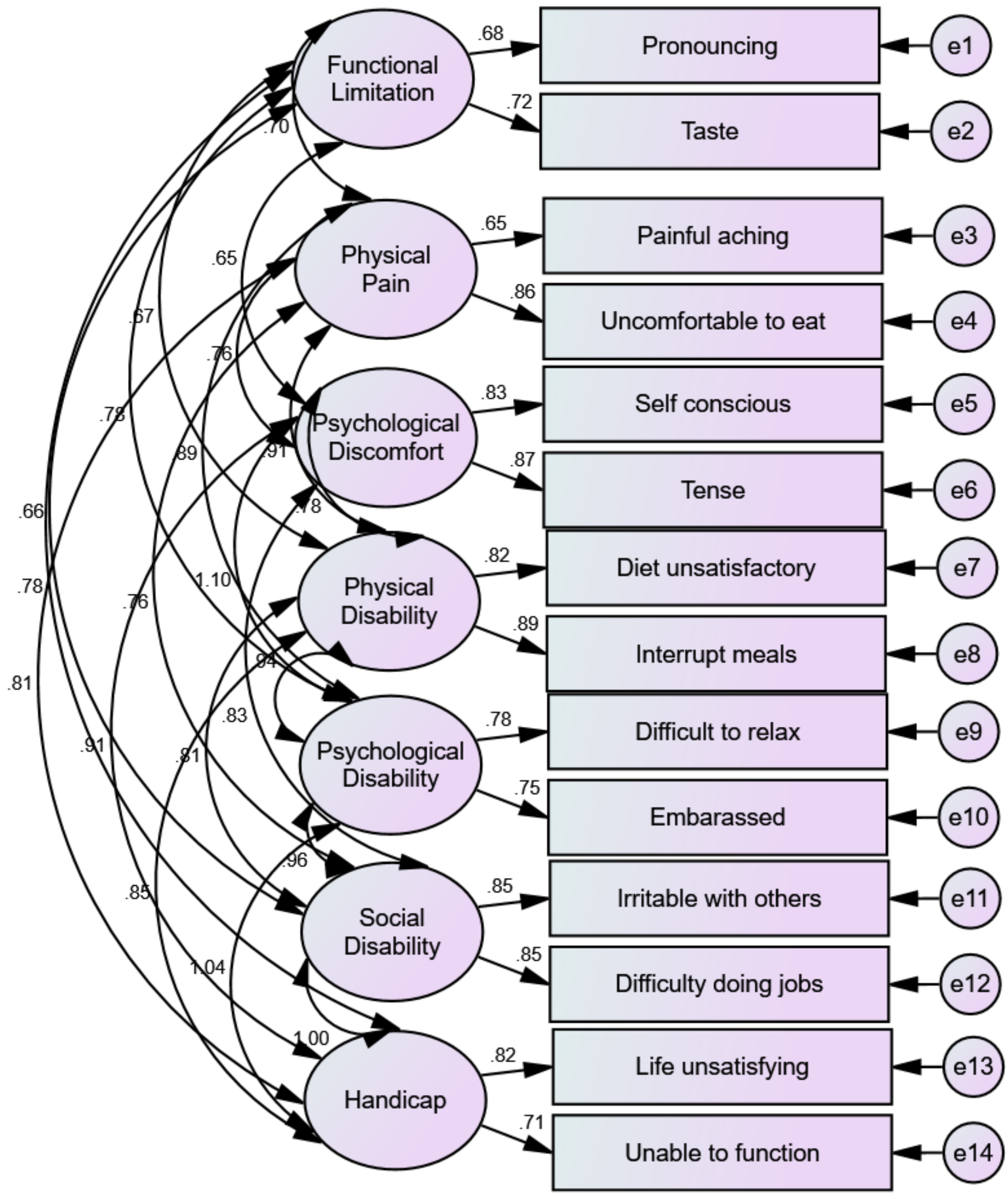

Figure 6.4 Seven factor model with estimates

The standardised regression weights for all the 14 items in four different models are given in Table 6.4. In the one factor model, the lowest regression weight was noted for item 1 which is 'difficulty in pronouncing words' and the highest was for item 9 which is 'difficulty to relax'. In the three factor model, the regression weights ranged from 0.62 to 0.88. In the six factor model, it ranged from 0.64 to 0.88 while in the seven factor model it was from 0.65 to 0.90.

Table 6.4 Standardised Regression weights for different models.

OHIP-14 Item	One Factor	3 Factor	6 Factor	7 Factor
1	0.50	0.67	0.67	0.68
2	0.53	0.73	0.73	0.73
3	0.61	0.62	0.64	0.65
4	0.74	0.80	0.86	0.86
5	0.73	0.80	0.80	0.83
6	0.84	0.88	0.81	0.87
7	0.76	0.75	0.90	0.82
8	0.80	0.85	0.85	0.90
9	0.87	0.85	0.85	0.78
10	0.74	0.76	0.88	0.75
11	0.80	0.81	0.84	0.85
12	0.78	0.78	0.80	0.85
13	0.82	0.83	0.82	0.82
14	0.70	0.69	0.71	0.71

The estimated variances of each item in different models are given in Table 6.5. The maximum variance estimated was observed for item 3 in all the four models and it was 1.47. The lowest variance for the one factor model was observed for item 9 (0.60) whereas in the three, six and seven factor models it was in item 8 and the variances were 0.48, 0.42 and 0.44 respectively.

Table 6.5 Variance estimates for different models

OHIP-14 Item	One Factor	Three Factor	Six Factor	Seven Factor
1	1.21	0.88	0.90	0.88
2	1.20	0.79	0.78	0.80
3	1.47	1.44	1.37	1.37
4	1.00	0.78	0.57	0.57
5	1.23	1.15	0.95	0.85
6	0.70	0.64	0.51	0.56
7	0.80	0.67	0.64	0.63
8	0.80	0.48	0.42	0.44
9	0.60	0.65	0.71	0.95
10	1.20	1.11	0.96	1.16
11	0.80	0.76	0.63	0.63
12	0.80	0.79	0.58	0.58
13	0.84	0.79	0.83	0.83
14	1.16	1.17	1.12	1.12

Table 6.6 Fit Indices for various models

Model	Chi-square	df	P value	χ^2/df	RMSEA (90% CI)	CFI	TLI	NFI	CMIN/df
1 Factor	699.51	77	<0.01	9.09	0.15 (0.14, 0.16)	0.83	0.80	0.82	9.09
3 Factor	562.33	74	<0.01	7.60	0.14 (0.13, 0.15)	0.87	0.84	0.85	7.60
6 Factor	472.98	62	<0.01	7.63	0.14 (0.13, 0.15)	0.89	0.84	0.88	7.63
7 Factor	410.07	56	<0.01	7.32	0.13 (0.12, 0.15)	0.90	0.84	0.89	7.32

Df = degrees of freedom; RMSEA = root mean square error of approximation; CFI = Comparative Fit Index; TLI = Tucker Lewis Index; NFI= Normed Fit Index

6.4 Discussion

This study mainly focussed on exploring the number of dimensions of OHIP-14 by examining different models that have been suggested by researchers using OHIP-14 data. As Confirmatory Factor Analysis is based on *a priori* specification of a model and OHIP is derived from Locker’s theoretical model, this was used in this study to explore the construct validity of various models that are in use. The findings of this study suggested that the current data failed to fit any of the four models used (one factor, three factor, six factor and seven factor models) and hence did not provide adequate construct validity. As the data failed to fit the single factor model, the evidence supports the view that the OHRQoL is a multidimensional construct. But, the other three models (three factor, six factor and seven factor), based on the multi-dimensional construct also did not fit the study data indicating these multi-dimensional models either poorly represent Locker's theoretical construct or the number of items considered may not sufficiently represent the OHIP. Hence as suggested by Baker and co-workers, further research on the items representing the Oral Health Related Quality of Life is needed. In all the models some of the factor correlations were very high, indicating that the factors are not discrete. As an example, in the three factor model, the correlation between the factors Pain-Discomfort (items 3 and 4 representing painful aching and uncomfortable to eat respectively) and Psychosocial impact (items 5,6,9,10 representing tense, feel conscious, embarrassed and difficult to relax respectively in OHIP-14) was 0.87 indicating that these two factors are similar and the existence of two separate factors for these items is not supported. As the correlations between some of the factors in the seven factor model exceeded 1 (Psychological Discomfort – Psychological Disability (1.10); Psychosocial Disability – Handicap (1.04); and Social Disability – Handicap (1.00), the model ended with Heywood's case. This is the situation where the correlation between two of the factors exceed the maximum value of 1. Though many reasons such as outliers (Bollen, 1987), non-convergence and under identification (Boomsma and Hoogland, 2001, Van Driel, 1978) and structurally mis-specified models (Bollen, 1989, Dillon et al., 1987, Kolenikov and Bollen, 2012, Sato, 1987, Van Driel, 1978) have been suggested for this situation, the last one, that is structural mis-specification, is of key importance. Hence, this analysis showed a lack of evidence for the existence of seven factors in OHIP-14 items.

Though varying opinions such as one factor (Santos et al., 2013), three factors (Montero et al., 2010), seven factors (Slade and Spencer, 1994) have been expressed regarding OHIP-14, the current data did not provide any evidence for the suitability of these models in OHIP-14. The present findings are of great importance in OHRQoL research as OHIP-14 is commonly used by researchers and clinicians to assess the influence of oral problems on day to day activities and to measure treatment effects. In the absence of adequate evidence for the construct validity of a model, the conclusions derived from the analysis based on these models are not reliable and great care should be taken when interpreting the findings (Santos et al., 2013).

7 Responsiveness of OHIP-14

7.1 Chapter Introduction

This chapter addresses the objective of measuring the responsiveness and classifying the patients as “Improved”, “No Change” and “Not Improved”. The responsiveness of the OHIP measure to change is important as this can be used to monitor longitudinal changes including treatment effect. The sensitivity of this measure to be able to assess the change in Oral Health Related Quality of Life over time has to be ensured in the process of measuring OHRQoL efficiently. As this study measured the OHRQoL of patients at three time points, this chapter addresses the issue of measuring the change scores and the sensitivity of the instrument for measuring change. In this chapter the responsiveness of the OHIP measure to assess the change in oral health related quality of life is explored. To do this Reliable Change (RC), Reliable Change Index (RCI) and Standard Error of Measurement (SEoM) are calculated and individual scores are estimated. Based on these measurements, the change in total OHIP-14 score over the period on each individual (change of state) was classified as “No change”, “Worsened” and “Improved” categories. The mean OHIP score for these three categories were also tested.

7.2 Methods

7.2.1 Conventional method

The change in Oral Health Related Quality of Life of patients measured using OHIP-14 has so far been addressed by comparing the mean scores and effect sizes. Although these conventional measures help to measure the overall change to some precision, they fail to address the individual changes. As treatment planning depends mainly on individual variations of a particular treatment, it is inevitable to address the individual variation in the effect of treatment for a dental condition. Moreover, the conventional methods give results with better precision when the sample sizes are reasonably large. The conventional methods of measuring minimally important differences in measuring Oral Health Related Quality of Life are effect size and Standard Error of Measurement (Revicki et al., 2008, Tsakos et al., 2012, Wyrwich et al., 2013). However, different opinions prevail between researchers in using these methods. Therefore, there is a necessity to compare these methods and to explore the possibility of better methods to measure the responsiveness of OHIP in dental research.

7.2.2 Reliability of change scores

There are several measures available in the literature to measure the change and its reliability. According to Lord and Novick (Lord et al., 1968), the reliability of change score is the ability of an instrument to detect the individual differences in change scores. This can be defined as:

$$\text{Reliability of change} = R_D = \sigma^2_D / (\sigma^2_D + \sigma^2_{err(D)})$$

where,

- σ^2_D is the systematic difference between subjects in their change score and,

- $\sigma^2_{err(D)}$ is the error in the estimate.

Based on the pre and post test scores, this can be expressed as follow.

$$\text{Reliability (D)} = R_D = \frac{\sigma^2_x R_{XX} + \sigma^2_Y R_{YY} - 2 * \sigma_X * \sigma_Y * r_{xy}}{\sigma^2_x + \sigma^2_Y - 2 * \sigma_X * \sigma_Y * r_{xy}}$$

where,

σ^2_X and σ^2_Y are the variance of pre and post test scores respectively

R_{XX} and R_{YY} are the reliability of pre and post-test measurements and,

r_{xy} is the correlation between pre-test and post test scores.

7.2.3 Reliable Change Index (RCI)

A method for measuring the Reliable and Clinically Significant Change (RCSC) has been proposed by Jacobson and co-workers (Jacobson et al., 1984) which measures the change at individual level in relation to the overall change based on the whole sample. The reliable change can be defined as the change in individual scores which fall above the change that is due to measurement variability. The change due to measurement variability is termed as Reliable Change Index (RCI) which is a measure of variation in Standard Error of Measurement (SEoM). This can be calculated using the following formula.

$$SEoM = SE_{diff} = SD_b \times \sqrt{1-R}$$

where,

- SD_b is the standard deviation of baseline observation

- and R is the reliability of the instrument which can be replaced with Cronbach's alpha.

Therefore,

$$RCI = (m_1-m_2) / SE_{diff}$$

where,

- m_1 is the subject's pre-test score and m_2 is the same subjects post test score.

The individual change scores can be measured using the above RCI. Further revision of this index has been carried out using the formula below (Hageman and Arrindell, 1993). This method has the advantage of predicting the actual change with more precision by addressing the problem of measurement unreliability which causes regression to the mean and can be applied to both negative and positive correlation between pre and post test scores. Also, this method is likely to protect from Type II errors.

$$RCI = \frac{(x_1 - X_2)\, r_{DD} + (m_1 - m_2)(1 - r_{DD})}{\sqrt{SEoM_1^{\,2} + SEoM_2^{\,2}}}$$

where,

RCI = Reliable Change Index

X_1 = Pre-treatment score

X_2 = Post-treatment score

m_1 = Pre-treatment mean score

m_2 = Post-treatment mean score

$SEoM_1$ = Standard Error of Measurement of Pre-treatment scores

$SEoM_2$ = Standard Error of Measurement of Post-treatment scores

r_{DD} = Reliability of difference scores.

Then the patients are classified into three categories according to their RCI scores (Eisen et al., 2007). The classification of patients into one of the three groups is based on the following rule:

1. Reliable decline (Worsened) if RCI < -1.96
2. No reliable change if -1.96 < RCI < 1.96 and
3. Reliable improvement if RCI > 1.96).

A similar classification was carried out using SEoM (Eisen et al., 2007):

1. Worsened if the post score increased more than 1 SEoM,
2. No change if the absolute value of the difference score is less than 1 SEoM and
3. Improved if the post score decreased more than 1 SEoM.

The responsiveness of OHIP from baseline to follow-up 2 was further compared using the paired t test for the three groups (classified as Worsened, No change and Improved using RCI) separately.

7.2.4 Simulated data for comparison

In order to validate the findings, a data set with 10000 samples were simulated using age, gender and total OHIP scores at three time points of the study. The simulation was conducted using SPSS. In the simulation command in SPSS the variables that were to be simulated were entered and the details about the distribution of each variable, minimum and the maximum possible values were also given. The simulation was carried out by assuming normal distribution for OHIP total scores at baseline, first and second follow-ups with a minimum of 0 (minimum total score based on all 14 items), and a maximum of 70 (maximum total score based on all 14 items) with the mean and standard deviation of the original values at each time point. Gender was assigned to follow Bernoulli distribution with values 0 and 1 whereas, age was based on the assumption of normality with the original mean and sd values based on the study data (mean age of 45.76 years with the standard deviation of 15.10 years). The simulated data were then stored in a separate file. Data for all the three time points (Baseline, follow-up 1 and follow-up 2) were generated and were used to calculate the different responsiveness index. The results obtained from study data and simulated data were then compared.

7.3 Results

7.3.1 Descriptive Statistics

The mean and standard deviation of the total OHIP scores for the three time points along with Cronbach's alpha value for the study data and simulated data are summarised in Table 7.2

Table 7.1 Mean and SD of composite scores at three time points.

Statistics	Time-0	Time-1	Time-2
Original Data			
Mean	21.63	20.53	18.49
Standard Deviation	15.75	16.51	16.64
Cronbach's alpha	0.94	0.95	0.96
Simulated Data			
Mean	18.20	17.71	16.64
Standard Deviation	14.44	15.20	14.35

The mean score at baseline was 21.63 out of the total of 70 and it gradually decreased at time points 1 and 2 in the original dataset. Though the mean values were slightly lower in the simulated data, a similar trend as in study data (decreasing mean values in the subsequent follow–ups) was found in the simulated data. The reliability of the instrument measured using Cronbach's alpha were 0.94, 0.95 and 0.96 for time points 0, 1 and 2 respectively.

7.3.2 Aggregate change

The overall change in Oral Health Related Quality of Life measured using OHIP-14 (responsiveness of OHIP-14) measured as the reliability of change using four different methods at three time points using study data and also based on the simulated data are given in Table 7.2. The change scores are classified as change occurring from baseline to 4 months after treatment (Time 0 and Time 2), change occurring from baseline to 2 months after treatment (Time 0 and Time 1) and change occurring from 2 months to 4 months after treatment (Time1 and Time 2). The classification between different time points is carried out to identify the actual change at these intervals.

Table 7.2 Change in OHIP score between Follow-up points

Measure	Time 0 and Time 2		Time 0 and Time 1		Time 1 and Time 2	
	Original	Simulated	Original	Simulated	Original	Simulated
RC	0.77	0.79	0.79	0.83	0.56	0.71
SEoM	3.79	3.48	3.66	3.36	3.29	2.83
Effect Size	0.20	0.04	0.07	0.11	0.12	0.07
RCI	0.83	0.47	0.29	0.17	0.56	0.32

RC = Reliability of change SEoM = Standard Error of Measurement

RCI = Reliability of Change Index

The reliability of change was 0.77 for the change scores from time0 and time2 and 0.79 for the change score from time0 and time1. The reliability of change was slightly smaller (0.56) for the change score from time1 and time2. Overall the reliability of change was much higher than 0 which demonstrates the ability of the instrument to measure the change in measuring Oral Health Related Quality of Life. The Overall indexes for measuring change based on all the individuals who provided data are given in terms of effect size, Standard Error of Measurement of change scores (SEoM) and Reliable Change Index. The RCI values are higher from baseline to second follow-up and it is least for baseline to first follow-up. A similar result was observed for effect size values. However, the SEoM was higher for baseline to second follow-up followed by baseline to first follow-up and least for the first to second follow-up. All the three measures showed that the change is maximum from baseline to second follow-up. Similar trend in responsiveness to change was observed for simulated data.

7.3.3 Individual Change scores

The individual change scores were assessed and all the patients were classified into three groups namely 'Worsened', 'No change' and 'Improved', based on the measures Reliable Change Index (RCI) and Standard Error of Measurement (SEoM) separately as explained in the methodology. The classification results are given in Table 7.3.

Table 7.3 Classification of patients using Responsiveness measures

Groups	Time0 Vs Time2 N (%)		Time0 Vs Time1 N (%)		Time1 Vs Time2 N (%)	
	Original	Simulated	Original	Simulated	Original	Simulated
Based on RCI						
Worsened	3 (4.00)	802 (8.02)	17 (19.10)	1785 (17.85)	2 (3.45)	1372 (13.70)
No Change	62 (82.67)	7827 (78.27)	57 (64.04)	5975 (59.75)	47 (81.03)	6388 (63.88)
Improved	10 (13.33)	1371 (13.71)	15 (16.85)	2240 (22.40)	9 (15.52)	2240 (22.40)
Based on SEoM						
Worsened	22 (29.33)	2495 (24.95)	27 (30.34)	2817 (28.17)	17 (29.31)	2466 (24.66)
No Change	30 (40.00)	3690 (36.9)	39 (43.82)	3668 (36.68)	23 (39.66)	4148 (41.48)
Improved	23 (30.67)	3815 (38.15)	23 (25.84)	3515 (35.15)	18 (31.03)	3386 (33.86)

Based on the revised RCI measure, 13.33% of the patients showed significant improvement in the OHRQoL at 4 months following initial treatment. A significant improvement after 2 months of initial treatment was noted in 16.85% of patients while 15.52% of the patients showed improvement between 2 months and 4 months after initial treatment. Only a small percentage (4.00%) of patients showed that their disease worsened during the period. A large percentage of patients (82.67%) recorded no change in Oral Health Related Quality of Life as measured by OHIP-14 during that period. The percentage of patients who significantly improved was higher when the difference scores are classified using Standard Error of Measurement (SEoM). The percentage of cases that worsened at all three time points were slightly higher in the simulated data set when compared to the original data set. Overall, the classification results based on RCI and SEoM using simulated data set were similar to the one obtained from the original data set.

The percentage of people who showed significant improvement was higher when the difference scores are classified using Standard Error of Measurement. The mean change in OHIP score from baseline to second follow-up was statistically significant for 'worsened' ($p=0.017$) and 'Improved' groups ($p<0.01$). However, the mean scores were not statistically significant for the 'No change' group ($p>0.05$).

Table 7.4 Summary statistics of baseline scores for the three groups classified using RCI and SEoM based on the change scores from baseline to follow-up-2.

Group	N	Mean (SD)	Median (Min, Max)
Based on RCI			
Worsened	3	27.67 (16.5)	28.00 (11, 44)
No Change	62	15.79 (15.41)	9.00 (0, 61)
Improved	10	40.20 (9.93)	39.00 (26, 57)
Based on SEoM			
Worsened	22	16.55 (15.77)	10.50 (0, 53)
No Change	30	14.07 (15.93)	8.50 (1, 61)
Improved	23	29.48 (15.54)	28.50 (6, 57)

The mean and standard deviation along with median of the total OHIP scores for the groups classified using RCI and SEoM are given in Table 7.4. There was a significant difference ($p<0.01$) in baseline scores between the three groups classified using RCI. Further post hoc analysis using Bonferroni correction revealed that the mean baseline score for the 'Improved' group was significantly higher than the 'No change' group ($p < 0.01$). The analysis carried on the groups classified using SEoM values also showed similar results. The mean baseline score differed significantly between the three groups ($p <0.01$). The post hoc analysis showed that the mean baseline score for the 'Improved' group was significantly higher than the 'Worsened' and 'No Change' groups ($p = 0.02$ and <0.01 respectively).

7.4 Discussion

Longitudinal changes in Oral Health Related Quality of Life scores have been studied by many researchers (Allen et al., 2001, Locker et al., 2004, Wong et al., 2007) using different methods. Locker and co-workers (Locker et al., 2004) using effect size and Guyatt Responsiveness Index reported that 6.2% of the samples showed a worsening of the condition. The present study using RCI has shown that only 4.00% reported a worsening condition. Locker's study found no significant difference between pre and post treatment scores for 'Improvement' group, but the current study showed a significant difference between pre and post treatment scores for patients who self-perceived as 'Improved' using RCI. A similar result was observed for the 'Worsened' group. As this is based only on 3 samples (only a small percentage self-perceived as worsened), interpretation of this finding should be viewed cautiously. However, the 'No change' group showed no significant difference, which corroborates Locker's findings.

The use of effect size has been criticized (Ferguson et al., 2002) as the use of standard deviation which is sample dependent and hence the results cannot be generalized. This is clear from the currently analysis that the effect size is very much affected in the simulated data set as the sample size is very large in the simulated data when compared to the original data. A related problem with current RC analyses is that the normal variation represented in the denominator is termed the Standard Error of the difference (SE_{diff}) (Liang, 2000), despite the fact that it is computationally the standard deviation of the individual scores at one point in time. A true estimate of change requires the standard deviation of difference scores (SD_{diff}) in the denominator. The suggested methods (SEoM and RCI) do not change from sample to sample and hence the results could be generalised. Also these methods have the potential to measure the change in the individual level.

This study examined and analysed the change in total OHIP score in broad spectrum. The change score depends on various parameters. As the main aim of this study was to explore the reliability of the change scores and to assess the number of individuals improved using the two methods (RCI and SEoM), other factors were not considered. It is clear from the analysis that the change is associated with the baseline scores. Patients who have higher OHIP total scores at baseline are more likely to show improvement during the follow-up. Also, the reasons for the attendance of patients at follow-ups vary considerably between different subjects mainly based on their severity of illness and the initial condition.

Though it has been established from the analysis that the oral health quality of life improved for the study patients, it could not be confirmed whether it is due to treatment, as no control group was involved in this study. Allen and co-workers (Allen et al., 2001) have also reported that there was a significant increase in OHRQoL after treatment and the original OHIP instrument with 49 items had good sensitivity to change. However, in this study, only a small percentage showed ‘Improvement’ during this period. This could be viewed from two angles. One is that the study period of 4 months may not be sufficient to assess the change in oral health related quality of life. As the majority of the patients attending King's College Dental clinic are 'Referral' patients, the nature of their illness is considered either severe or requiring extensive or lengthy courses of treatment. In this case, it may not be possible to see a noticeable change within four months of initial treatment. In that case, a longer follow-up period is suggested. The second one could be that the instrument used to measure (OHIP-14) may not be efficient enough to measure the change. But this option is ruled out as this study has revealed that OHIP-14 has good reliability and also has the reliability to measure the change in oral health. It may be concluded that to study the full change in oral health in this subject group, more follow-up periods are necessary.

8 Validation of methodology with a large independent dataset

8.1 Chapter introduction

In this chapter, the methods suggested for handling various methodological issues in OHIP-14 were tested using a large nationwide dataset. The dataset used for this validation was the Adult Dental Health Survey Data 2009 and consisted of 11,380 individuals. In this data only 30 cases (0.26%) had missing data, so the missing data were simulated and the simulated data were used for testing the missing data handling techniques. These data were used for all the methods suggested and the results were compared with the results obtained from the research sample.

8.2 Description of National Sample

The Adult Dental Health Survey (ADHS) (UKDA study number 6,884) is a series of national dental surveys that have been carried out in England and Wales (1968 and 1978), England, Wales, Scotland and Northern Ireland (1988 and 1998) to study the oral health of the adult population and how it has changed since 1968. The latest survey was carried out in 2009 (delayed by a year) in England, Wales and Northern Ireland. It was commissioned by NHS Information Centre for Health and Social Care (NHS IC). Scotland did not participate in the 2009 survey. The data collection took place between October 2009 and April 2010 and was first published as a National Statistic on 24 March 2011. The study details including methodology, the abbreviations used and the measures adopted can be found in O’Sullivan (O'Sullivan et al., 2011). The survey was managed by Office of the National Statistics, Social Survey Division Information Centre for Health and Social Care and a total of 11,380 individuals were interviewed. The data comprised of demographic, oral health, clinical and behavioural details of all the individuals (Foundation report: Adult Dental Health Survey 2009). As 30 samples contained missing values, they were removed and the remaining 11350 samples were considered as the comparison data with the study sample.

8.3 Validation of Methodology

8.3.1 Missing Data in Reference data set

As there were no missing data in the Adult Dental Health Survey data set sample, missing data were created in order to test the suitability and efficiency of different missing data handling methods for using in OHIP data. As the missing pattern in the study sample was missing completely at random, a similar situation was simulated by selecting 2838 subjects (25.00% of the total sample) randomly to have missing OHIP scores. This was done by generating 2838 random numbers using the Excel function “RANDBETWEEN”, where the generated numbers represent the participant numbers whose total OHIP scores were removed and treated as the missing data. In order to confirm the randomness of missingness in the data, the mean total OHIP scores were compared between missing and non-missing cases. There was no significant difference between missing and non-missing cases with respect to mean total OHIP scores (p=0.68). It was also shown that there was no significant association between missingness and other demographic variables. The method of logistic regression was used to prove that none of the demographic variables predicted the missingness in the data. In this process, cases with missing scores were coded as 1 and cases without missing cases were coded as 0. This binary variable was used as an outcome measure and demographic variables, namely age group, gender and ethnic groups were treated as predictor variables. So, the comparison data were prepared where in 8512 cases had complete data and 2838 had missing OHIP total scores, out of the total sample of 11350.

To make sure that the missing data generated in the comparison sample follow MCAR pattern, Little’s MCAR test was carried out and it confirmed that the missing data generated followed Missing Completely At Random pattern ($p=0.07$). To find out whether any of the demographic variables predicted the missingness, a logistic regression analysis was carried out using missingness as outcome measure and age, sex, ethnicity, education and profession as predictor variables. This showed that none of the variables predicted ($p>0.05$ for all variables) missingness. As none of the demographic variables predicted missingness and Little’s MCAR test was not significant, there is a strong evidence that the missing data generated in the comparison data set followed the Missing Completely At Random pattern.

As described in Chapter 6, a number of different methods tested to handle the missing data were adopted and these methods were compared. The missing data were filled with item mean, subject mean, interpolation, regression method, using trend in the data, Expectation and Maximization (EM) algorithm and Multiple Imputation methods. The data created with the above methods were analysed for comparison of the methods apart from complete case analysis. Mean, standard deviation and skewness were used to compare these methods. Table 8.1 summarises these measure for different methods of imputation of missing values.

Table 8.1 Estimates of Mean and sd values using different missing data techniques (Based on ADHS 2009 sample)

Method	Mean	Standard deviation	Median	Range	Skewness
Based on all data	**3.40**	**6.22**	**1.00**	**56**	**3.06**
Completed case	3.42	6.24	1.00	56	3.05
Item mean	3.40	5.76	1.29	56	3.17
Subject mean	3.41	6.29	0.00	56	3.05
Interpolation	3.40	5.84	1.00	56	3.07
Regression	3.42	6.18	1.00	58.12	3.06
Trend	3.40	5.76	1.29	56	3.17
EM Algorithm	3.40	6.17	0.88	56.22	3.08
Multiple Imputation	3.40	6.21	1.00	64.50	3.02

Comparison of methods

The mean, standard deviation, median, range and skewness calculated using eight missing data handling methods for the comparison data set are given in the above table. All the eight different methods showed similar mean values with the mean based on all data. However, the standard deviation and skewness differed between the methods. The EM Algorithm and Multiple Imputation methods showed similar values for all the measures except for range. Therefore, all the missing data handling methods used in theses analyses predicted similar values for the missing data.

The generated missing data samples were further analysed to find out which method gives better estimates for the mean total score. The 2838 cases for which the missing data were artificially simulated were considered for the analyses. The original total OHIP-14 scores of the selected 2838 cases were compared with the new values (imputed values) obtained by implementing different missing data handling methods. The mean difference between original score and imputed values using different methods along with sd, median, minimum and maximum are summarised in Table 8.2 for all the seven missing data methods. The mean imputed score was significantly different from the original mean for Regression and multiple imputation methods. The rest of the methods showed no significant difference ($p>0.05$ for all the methods) from the original values.

Table 8.2 Summary statistics of difference between original and imputed values

Method	Mean (sd)	Median (min, max)	P value
Original	3.36 (6.18)	1.00 (0, 52)	
Item mean	3.34 (4.03)	1.90 (0, 37)	0.66
Subject mean	3.38 (6.45)	0 (0, 56)	0.71
Regression	3.55 (5.94)	1.44 (0, 54)	<0.01
Trend	3.34 (4.03)	1.90 (0, 37)	0.67
Interpolation	3.35 (4.43)	2.00 (0, 36)	0.85
EM Algorithm	3.34 (5.96)	0.75 (0, 53)	0.50
Multiple Imputation	3.48 (5.97)	0.97 (0, 52)	<0.01

The closeness between original and imputed values were assessed using Bland-Altman statistics and graphs. The Bland-Altman statistics (based on original and imputed data) for the seven missing data handling methods are summarised in Table 8.3. The mean difference between original and simulated values was close to 0 for all the methods except regression and multiple imputation methods. The standard deviation of the mean difference was lowest for EM algorithm.

Table 8.3 Bland - Altman statistics for imputed data from different methods

Method	Mean Difference (sd)	Mean $\pm$2sd Limits	Standard Error	P value	No. outside 2sd limits (%)
Item mean	0.02 (2.66)	-5.29 to 5.33	0.05	0.66	134 (4.72)
Subject mean	-0.01 (1.82)	-3.66 to 3.63	0.03	0.70	176 (6.20)
Regression	-0.07 (1.87)	-3.82 to 3.68	0.04	0.04*	160 (5.64)
Trend	0.02 (2.66)	-5.29 to 5.33	0.05	0.66	133 (4.69)
Interpolation	0.01 (3.19)	-6.36 to 6.38	0.06	0.85	173 (6.10)
EM Algorithm	0.02 (1.52)	-3.02 to 3.06	0.03	0.51	154 (5.43)
Multiple Imputation	-0.23 (1.58)	-3.40 to 2.94	0.03	0.01*	146 (5.14)

* denotes the mean difference is significantly different from 0.

Bland Altman graph for the seven methods of missing data handling methods are given in Figure 8.1 (a to d) and Figure 8.2 (e to g).

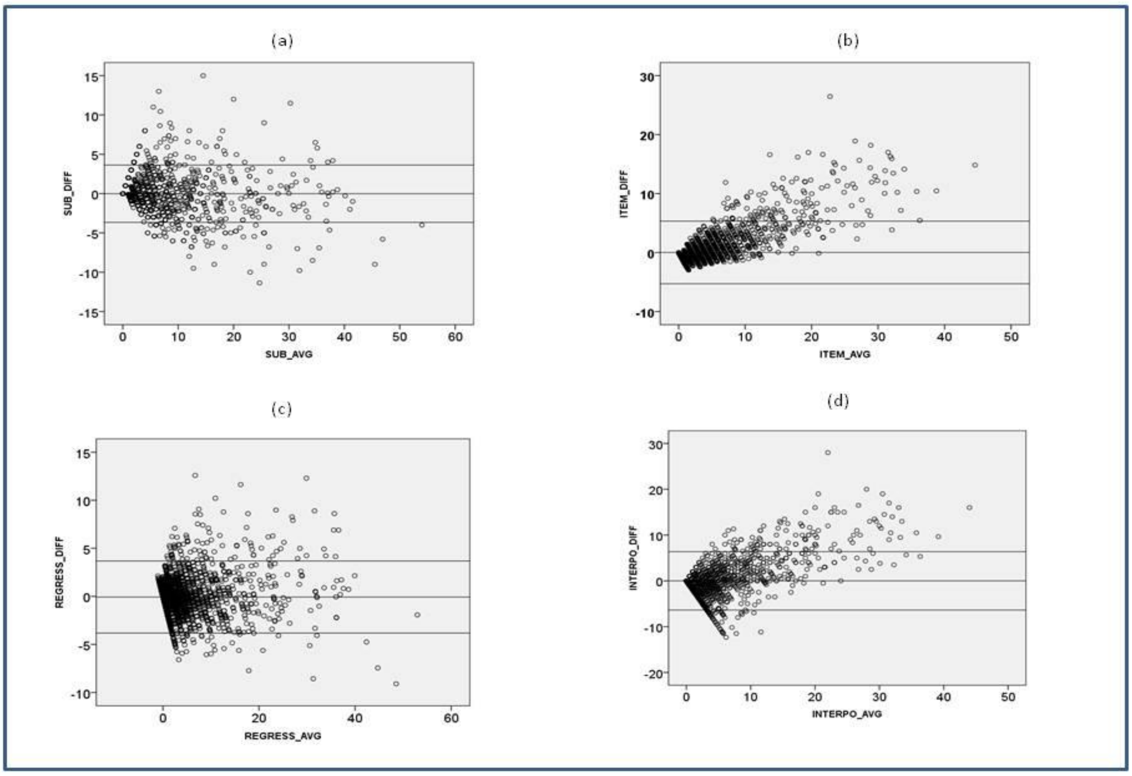

Figure 8.1 Bland Altman graphs for comparing the two measurements (Original and imputed total OHIP scores) for different missing data handling techniques.

The pictures showing the relationship between average (of original and imputed values) and the difference (between original and imputed values) using data obtained from a) subject average b) item average c) regression and d) interpolation imputation methods.

The horizontal lines represent the mean and mean <u>+</u> 2 sd of the difference.

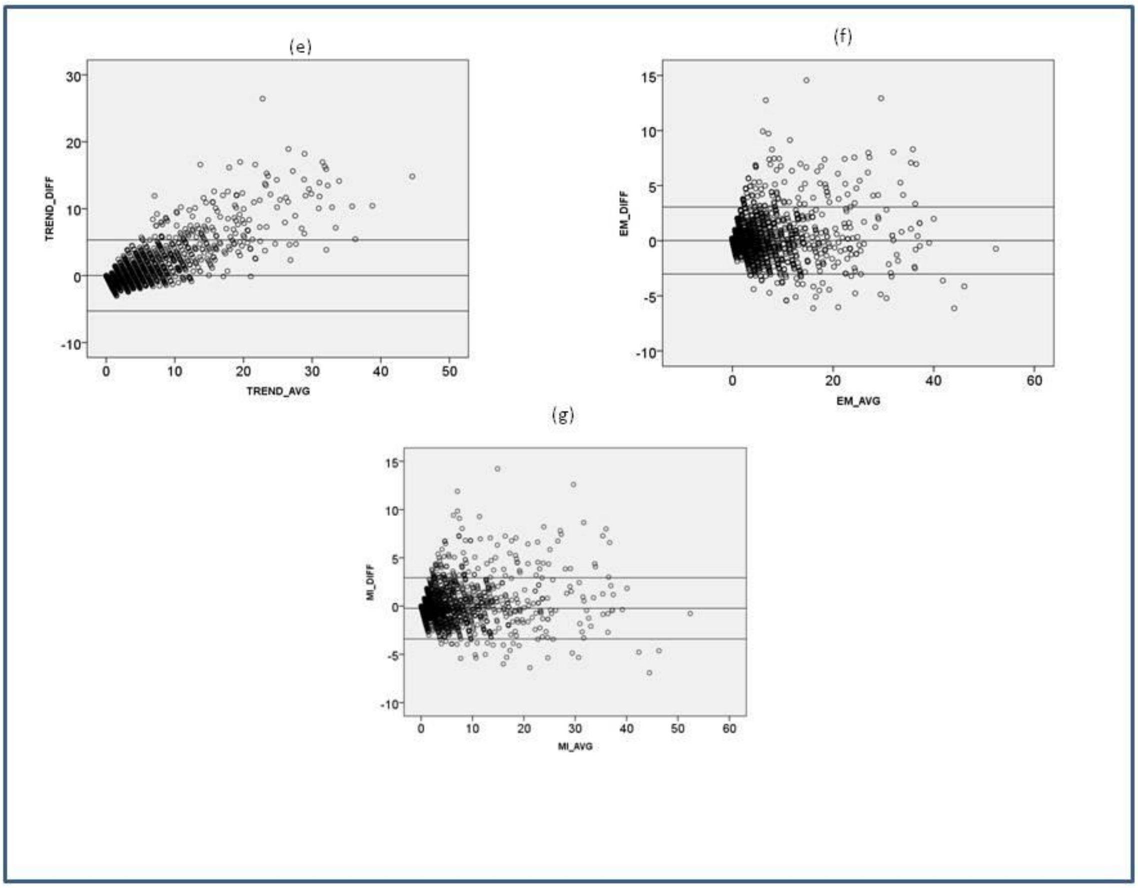

Figure 8.2 Bland Altman graphs for comparing the two measurements (Original and imputed) for different missing data handling techniques.

The pictures showing the relationship between average (of original and imputed values) and the difference (between original and imputed values) using data obtained from e) Trend f) EM algorithm and g) Multiple Imputation methods.

The horizontal lines represent the mean and mean $\pm$ 2 sd of the difference.

The Bland Altman graph for different methods showed a random variability for EM algorithm, Regression, Subject Mean and Multiple Imputation methods. An increase in variability was observed for the methods Item Mean, Interpolation, and Trend. The mean difference between original and imputed values shows the estimated bias and the standard deviation of the difference measures the random fluctuations. The significant p values for Regression and Multiple Imputation methods show that the mean difference is significantly different from 0 indicating the fixed bias in these methods.

8.3.2 Floor and Ceiling Effects

From the Adult Dental Health Survey 2009 data in which 11350 samples were considered, none had the OHIP item missing. The OHIP-14 was scaled from 0 to 4 and hence the total OHIP score (the composite score) for any individual ranged between 0 and 56. Both for individual items and for the total score, the higher scores indicate poorer OHRQoL of life whereas the lower score indicates better OHRQoL. The number of participants who gave the lowest score (floor) and the highest score (ceiling) for each of the 14 items are shown in Figure 8.3.

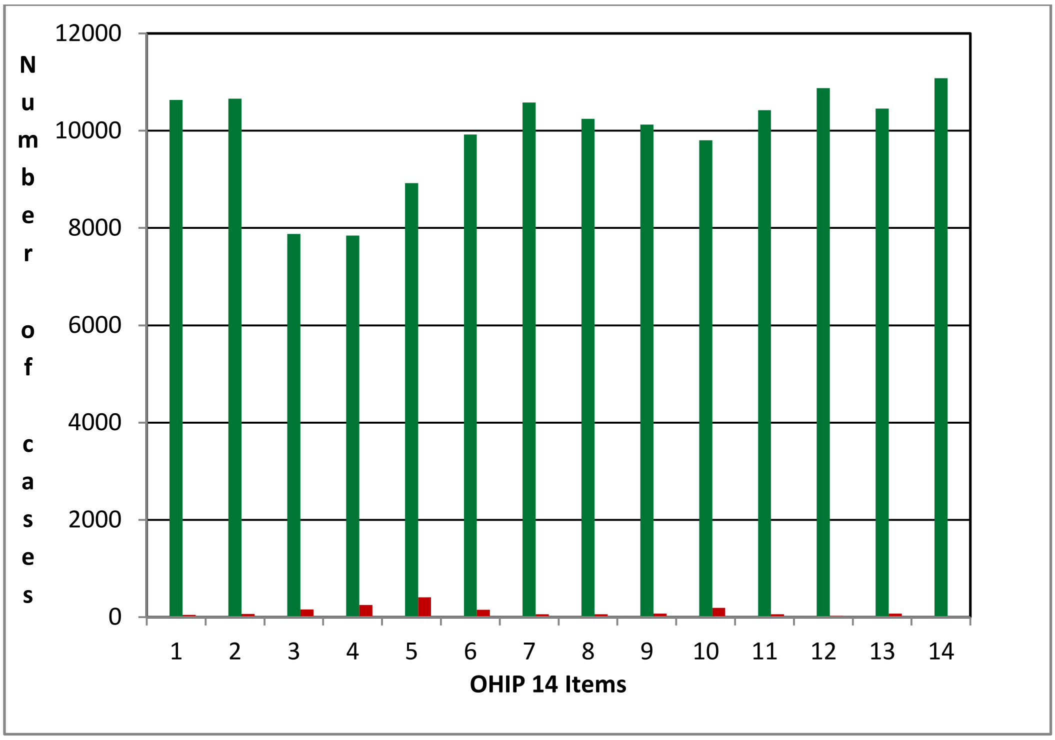

Figure 8.3 The bar represents the number of participants who have given lowest (0) (the green bar) and highest (5) scores (the red bar) for each of OHIP-14 items.

As in the study sample described previously, it is clear that the floor score is seen considerably more often than the ceiling score in the ADHS data. The amount of left censoring (floor) ranged between 69.10% and 97.60% for the 14 items, while the amount of right censoring (ceiling) ranged between 0.14% and 3.60%. The average item difficulty (average item sufferings) was 0.88 indicating that on average 88% of the participants consistently gave the lower end of the score. On the right censoring (ceiling), the average item easiness was 0.01, meaning that around 1.00% of the participants tended to give the higher side of the score. Therefore, based on the method suggested by Ven den Oord and Van den Ark (Oord and Ark, 1997), the censoring points at lower side (floor) and higher side (ceiling) for the composite score of OHIP-14 were 8.40 and 29.60 respectively.

Table 8.4 Results of OLS and Tobit models using ADHS 2009 data

Predictors	Ordinary Linear Regression (OLS)		Tobit Regression	
	Effect (95% CI)	**p value**	**Effect (95% CI)**	**p value**
TOTAL OHIP-14				
Age (in years)				
35 - 54	-0.07 (-0.40, 0.26)	0.67	0.06 (-1.32, 1.44)	0.93
55 - 74	-0.28 (-0.62, 0.06)	0.11	-1.09 (-2.53, 0.36)	0.14
75 and above	-1.00 (-1.45, -0.55)	0<0.01	-4.22 (-6.29, -2.15)	<0.01*
Gender	-0.43 (-0.68, -0.19)	<0.01	-0.43 (-0.68, -0.19)	<0.01*
Profession				
Managerial	-0.30 (-0.80, 0.20)	0.24	-0.30 (-0.80, 0.20)	0.24
Employee	0.05 (-0.33, 0.42)	0.81	0.05 (-0.33, 0.42)	0.81
FUNCTIONAL LIMITATION				
Age (in years)				
35 - 54	0.06 (0.02, 0.11)	<0.01	0.61 (0.21, 1.01)	<0.01*
55 - 74	0.17 (0.12, 0.22)	<0.01	1.35 (0.94, 1.75)	<0.01*
75 and above	0.18 (0.12, 0.24)	<0.01	1.42 (0.93, 1.91)	<0.01*
Gender	0.02 (-0.01, 0.06)	0.22	0.17 (-0.09, 0.43)	0.20
Profession				
Managerial	-0.03 (-0.10, 0.04)	0.37	-0.43 (-1.02, 0.15)	0.15
Employee	0.06 (0.01, 0.12)	0.02	-0.40 (-0.00, 0.81)	0.05
PHYSICAL PAIN				
Age (in years)				
35 - 54	-0.04 (-0.13, 0.06)	0.42	-0.07 (-0.27, 0.14)	0.54
55 - 74	-0.08 (-0.18, 0.02)	0.10	-0.13 (-0.35, 0.08)	0.23
75 and above	-0.31 (-0.44, -0.18)	<0.01	-0.65 (-0.94, -0.36)	<0.01*

Predictors	Ordinary Linear Regression (OLS)		Tobit Regression	
	Effect (95% CI)	p value	Effect (95% CI)	p value
Gender	-0.08 (-0.14, -0.01)	0.03	-0.20 (-0.36, -0.05)	0.01*
Profession				
Managerial	-0.04 (-0.18, 0.10)	0.57	-0.01 (-0.32, 0.31)	0.97
Employee	-0.02 (-0.12, 0.09)	0.73	0.04 (-0.20, 0.28)	0.74
PSYCHOLOGICAL DISCOMFORT				
Age (in years)				
35 - 54	0.07 (-0.02, 0.16)	0.12	0.28 (0.01, 0.55)	0.04*
55 - 74	-0.03 (-0.12, 0.06)	0.57	-0.01 (-0.29, 0.28)	0.97
75 and above	-0.31 (-0.43, -0.19)	<0.01	-1.07 (-1.48, -0.67)	<0.01*
Gender	-0.20 (-0.26, -0.14)	<0.01	-0.65 (-0.86, -0.45)	<0.01*
Profession				
Managerial	-0.10 (-0.23, 0.03)	0.13	-0.36 (-0.78, 0.07)	0.10
Employee	-0.004 (-0.09, 0.10)	0.93	-0.00 (-0.31, 0.31)	1.00
PHYSICAL DISBILITY				
Age (in years)				
35 - 54	-0.04 (-0.09, 0.01)	0.14	-0.19 (-0.58, 0.19)	0.32
55 - 74	-0.03 (-0.08, 0.03)	0.34	-0.13 (-0.52, 0.27)	0.52
75 and above	0.03 (-0.04, 0.10)	0.47	0.20 (-0.31, 0.71)	0.43
Gender	-0.03 (-0.07, 0.01)	0.16	-0.28 (-0.56, 0.01)	0.06
Profession				
Managerial	-0.04 (-0.12, 0.04)	0.35	-0.37 (-0.98, 0.24)	0.24
Employee	0.01 (-0.04, 0.07)	0.64	0.11 (-0.33, 0.54)	0.63

Predictors	Ordinary Linear Regression (OLS)		Tobit Regression	
	Effect (95% CI)	p value	Effect (95% CI)	p value
PSYCHOLOGICAL DISABILITY				
Age (in years)				
35 - 54	-0.05 (-0.12, 0.12)	0.15	-0.14 (-0.41, 0.12)	0.29
55 - 74	-0.14 (-0.21, -0.07)	<0.01	-0.52 (-0.80, -0.24)	<0.01*
75 and above	-0.29 (-0.39, -0.20)	<0.01	-1.27 (-1.68, -0.86)	<0.01*
Gender	-0.10 (-0.14, -0.05)	<0.01	-0.47 (-0.68, -0.27)	<0.01*
Profession				
Managerial	-0.03 (-0.14, 0.07)	0.50	-0.18 (-0.61, 0.26)	0.43
Employee	0.04 (-0.03, 0.12)	0.25	0.23 (-0.09, 0.55)	0.16
SOCIAL DISABILITY				
Age (in years)				
35 - 54	-0.05 (-0.09, -0.01)	0.03	-0.21 (-0.57, 0.16)	0.27
55 - 74	-0.14 (-0.18, -0.09)	<0.01	-1.00 (-1.41, -0.59)	<0.01*
75 and above	-0.23 (-0.29, -0.17)	<0.01	-2.45 (-3.16, -1.74)	<0.01*
Gender	0.0.04 (-0.07, -0.01)	0.01	-0.45 (-0.75, -0.15)	<0.01*
Profession				
Managerial	-0.02 (-0.09, 0.04)	0.45	-0.26 (-0.86, -0.34)	0.40
Employee	-0.03 (-0.08, 0.02)	0.24	-0.29 (-0.74, 0.15)	0.20
HANDICAP				
Age (in years)				
35 - 54	-0.01 (-0.05, 0.03)	0.48	0.06 (-0.34, 0.47)	0.77
55 - 74	-0.02 (-0.06, 0.02)	0.42	-0.01 (-0.43, 0.41)	0.97
75 and above	-0.06 (-0.11, -0.001)	<0.05	-0.36 (-0.94, 0.22)	0.22

Predictors	Ordinary Linear Regression (OLS)		Tobit Regression	
	Effect (95% CI)	**p value**	**Effect (95% CI)**	**p value**
Gender	-0.01 (-0.04, 0.02)	0.66	-0.08 (-0.38, 0.22)	0.59
Profession				
Managerial	-0.03 (-0.09, 0.03)	0.35	-0.24 (-0.86, 0.34)	0.44
Employee	-0.01 (-0.06, 0.03)	0.55	-0.14 (-0.59, 0.31)	0.53

Note: Comparison groups: age = 16 - 34 Gender = Male Profession = Self employed

*Denotes statistically significant.

The results of the ordinary least squares (OLS) and Tobit regressions are given in Table 8.4 for the total OHIP-14 and for each of the seven dimensions. Both OLS and Tobit models showed that none of the variables significantly (p>0.05) predicted the total OHIP scores. However, the Tobit model, after adjusting for floor and ceiling effects produced higher values for the regression coefficients indicating that the estimated effect of predictor variables on the total OHIP score is higher under the Tobit model than under the OLS regression model.

Further dimension wise analysis also showed similar results for all the seven dimensions. For the dimensions namely, Functional Limitation, Physical pain, Psychological discomfort, Physical disability and Handicap none of the predictor variables were statistically significant in either OLS or Tobit models. Participants of age 75 or more had significantly lower scores for psychological disability when compared to 16-34 age group, which was statistically significant (p<0.05) both in the OLS and Tobit models. Similarly, the score for psychological disability dimension was significantly lower (p<0.05) for participants who had not specified their profession when compared to self-employed people. The Tobit model also predicted gender as a significant predictor (p=0.04) of psychological disability whereas it was not significant in OLS. For the social disability dimension, both OLS and Tobit models identified age as a significant predictor. Participants with 75 or older had significantly lower scores when compared to the young age group (16-34).

As previously observed, the results of the significance tests for the regression coefficients were very consistent between the OLS and Tobit models. However, the estimate of the regression coefficients was much larger under Tobit regression, confirming that the OLS, without considering floor and ceiling effects, under-estimates the effects.

Model comparison

The suitability of both the regression and Tobit models for the ADHS OHIP data was explored by comparing log likelihood and the model fit statistics. These statistics for the model with total OHIP score and for the dimension scores are given in Table 8.5.

Table 8.5 Model fit statistics for OLS and Tobit models (For comparison data set)

Outcome	Model	P value	Log likelihood
Total OHIP Score	OLS Tobit	<0.01 <0.01	-33522.12 -7743.21
Functional Domain	OLS Tobit	<0.01 <0.01	-13122.26 -3704.72
Physical pain	OLS Tobit	<0.01 <0.01	-20567.52 -12166.52
Psychological discomfort	OLS Tobit	<0.01 <0.01	-19678.28 -8555.47
Physical disability	OLS Tobit	0.09 0.06	-14345.10 -4385.56
Psychological disability	OLS Tobit	<0.01 <0.01	-16985.48 -6461.70
Social disability	OLS Tobit	<0.01 <0.01	-12603.67 -3496.14
Handicap	OLS Tobit	0.53 0.76	-11717.74 -3065.29

The absolute log likelihood values for the Tobit model are much lower when compared to the OLS regression model for total OHIP score and for all the seven dimensions indicating that the Tobit model gives a better fit than OLS regression model. The p values of both OLS regression and Tobit models for the dimensions Physical disability and Handicap were greater than 0.05 indicating that the predictor variables (age, gender and profession) do not show a significant relationship with the OHIP scores of corresponding dimensions. However, the p values of both OLS regression and Tobit models for total OHIP score and other dimensions (Functional, Physical Pain, Psychological Discomfort, Psychological Disability and Social Disability) were statistically significant ($p<0.01$ for all) indicating that the predictor variables (age, gender and profession) show a significant relationship with the corresponding OHIP scores.

8.3.3 Factorial Structure

The factorial structure of OHIP items was earlier checked with the study data. In order to validate the findings, various models considered for the analyses (one factor, three factors, six factors and seven factors models) were tested using the reference data set. As in the study data, the confirmatory factor analyses were carried out separately based on each of the four models and the results are given in this section.

8.4 Results

The reference data set was used to confirm the number of factors in OHIP-14 items using four different models discussed above. The Confirmatory Factor Analyses carried out using the reference data set showed that none of the models fitted the data corroborating the findings using the study data set. The fitted models are shown in Figure 8.4 to Figure 8.7. The model fit was assessed using the same indices used in the study data and are shown in Table 8.8. Out of the seven criteria used to test the fit of the model, only two (CFI>= 0.90 and NFI>=0.90) met the criteria in the seven factor model whereas none of the criteria were met for the other three models. Further analysis of factor correlations revealed that the factors in these models are highly correlated questioning the discriminant validity of OHIP-14 items. In the three factor model, the correlation between Pain-Discomfort and Psychosocial Impact, exceeded the value of 0.85 (r=0.88). The correlation between four different pairs of factors namely - Physical Disability – Handicap (r=0.86), Psychological Impact – Social Disability (r=0.85), Psychological Impact – Handicap (r=0.87) and Social Disability – Handicap (r=0.91) exceed the value of 0.88 in the six factor model. Similarly, in the seven factor model, the correlation between six different pairs of factors exceed the value of 0.85 out of which one correlation exceeded 1 (Psychological discomfort – Psychological Disability) indicating the presence of a Heywood case. The other five pairs of factors that were highly correlated were: Physical Disability – Psychological Disability (r=0.96), Handicap (0.86); Psychological Disability – Social Disability (0.99), Handicap (0.99); Social Disability - Handicap (0.91). The findings indicate that the constructs in these models are not distinct and hence have poor discriminant validity.

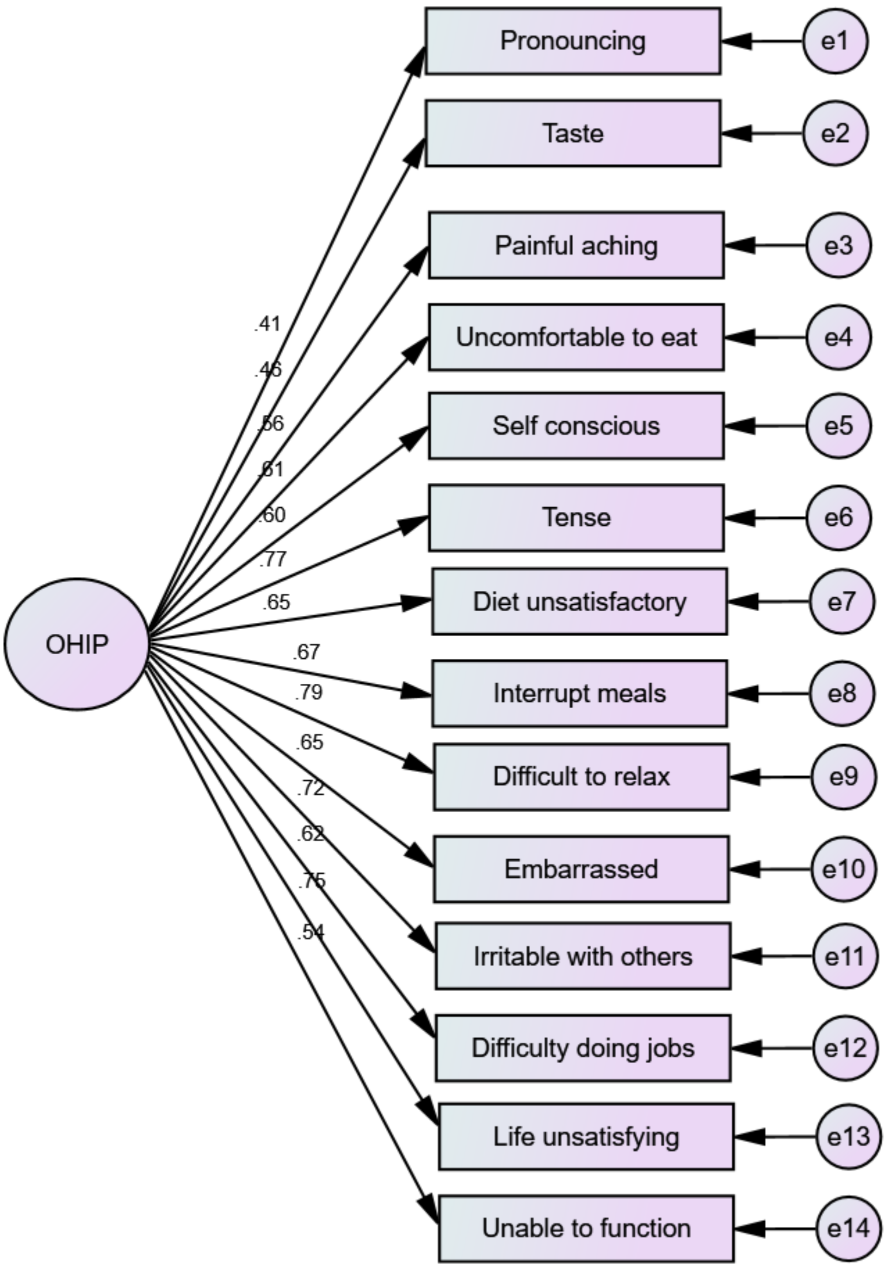

Figure 8.4 Fitted Single factor Structural Equation Model for OHIP-14 items depicting the relationship between exogenous variables (factors) and endogenous variables (OHIP items) along with covariances.

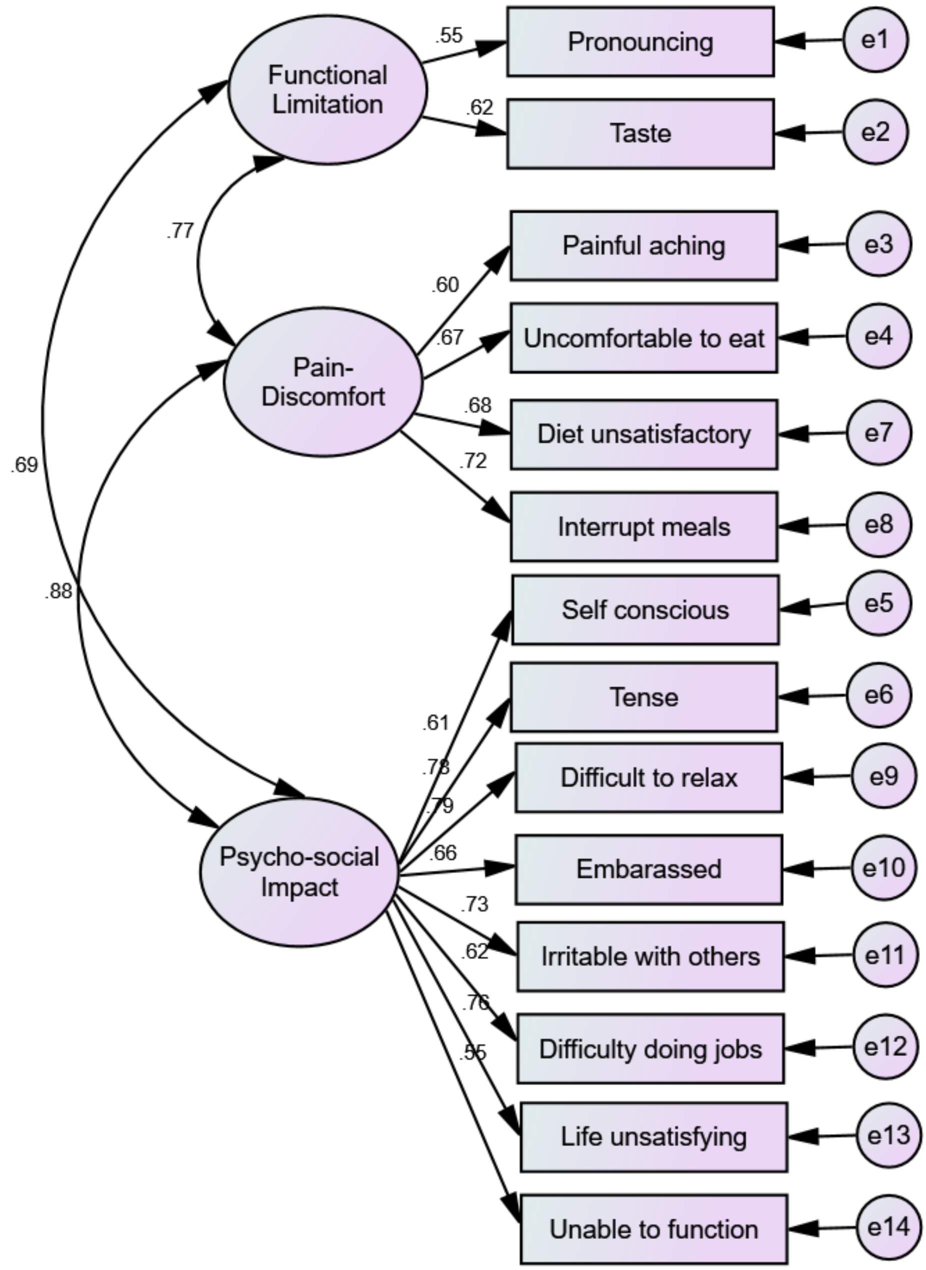

Figure 8.5 Three factor model with estimates

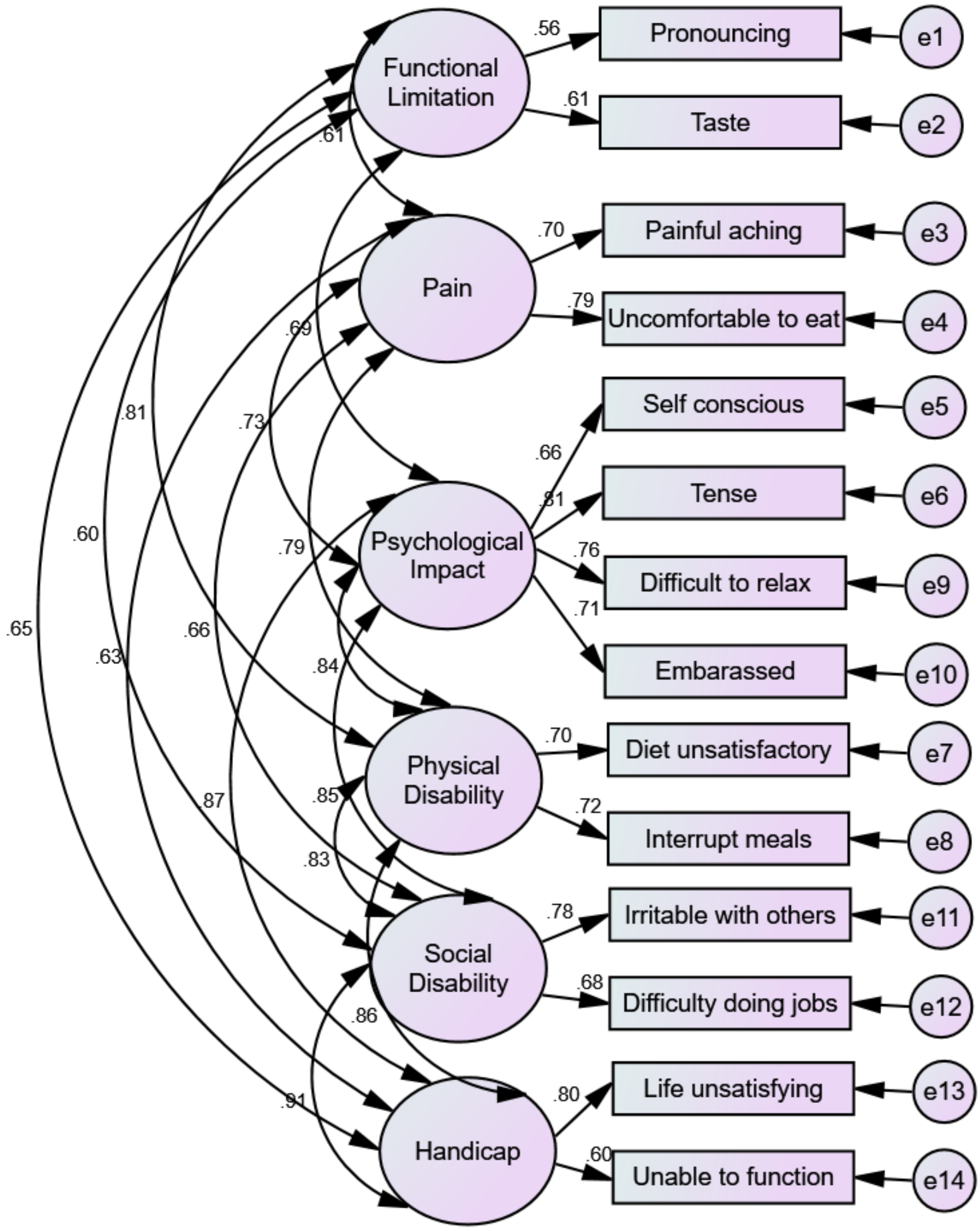

Figure 8.6 Six factor model with estimates

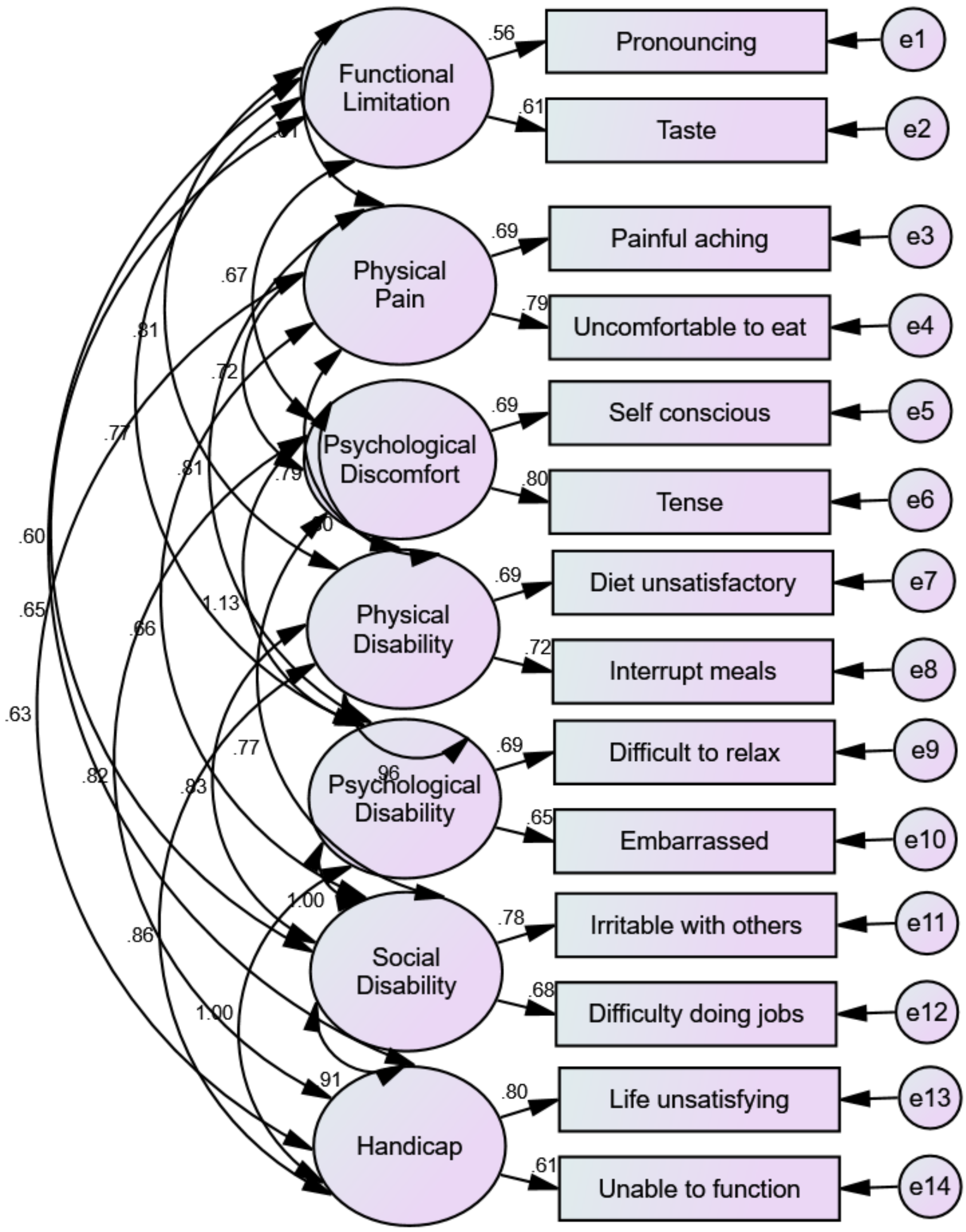

Figure 8.7 Seven factor model with estimates

Table 8.6 shows the standardised regression weights of all the 14 OHIP items based on the four different models fitted using reference data. As in the study data, the lowest regression weight was observed for item-1 in the one factor model while the highest was for item 9. The regression weights ranged between 0.55 to 0.79, 0.61 to 0.81 and 0.56 to 0.80 for three, six and seven factor models respectively. These findings were similar to those obtained using the study data.

Table 8.6 Standardised Regression weights for different models.

OHIP-14 Item	One Factor	3 Factor	6 Factor	7 Factor
1	0.41	0.55	0.56	0.56
2	0.46	0.62	0.61	0.61
3	0.56	0.60	0.70	0.69
4	0.61	0.67	0.79	0.79
5	0.60	0.68	0.66	0.69
6	0.77	0.72	0.70	0.80
7	0.65	0.61	0.72	0.70
8	0.67	0.78	0.78	0.72
9	0.79	0.79	0.68	0.69
10	0.65	0.67	0.81	0.65
11	0.72	0.73	0.77	0.78
12	0.62	0.62	0.71	0.68
13	0.75	0.76	0.80	0.80
14	0.54	0.55	0.61	0.61

The variance estimates of the OHIP-14 items based on the four different models using the reference data are shown in Table 8.7. The maximum estimated variance of 0.68 was observed for item 4 in one factor model whereas all the other three models showed the maximum values of 0.71, 0.63 and 0.58 respectively for item 5. All the four models recorded the minimum estimated variance of 0.06 for item 14.

Table 8.7 Variance estimates for different models

OHIP-14 Item	**One Factor**	**Three Factor**	**Six Factor**	**Seven Factor**
1	0.22	0.19	0.18	0.18
2	0.23	0.18	0.18	0.18
3	0.64	0.60	0.48	0.49
4	0.68	0.59	0.41	0.41
5	0.72	0.71	0.63	0.58
6	0.24	0.23	0.20	0.21
7	0.17	0.16	0.15	0.15
8	0.19	0.17	0.17	0.17
9	0.16	0.16	0.17	0.21
10	0.39	0.38	0.34	0.39
11	0.15	0.15	0.12	0.12
12	0.09	0.09	0.07	0.07
13	0.15	0.14	0.12	0.12
14	0.06	0.06	0.06	0.06

Table 8.8 Fit Indices for four different models.

Model	Chi-square	df	P value	χ^2/df	RMSEA (90% CI)	CFI	TLI	NFI	CMIN/df
1 Factor	11689.91	77	<0.01	151.82	0.12 (0.11, 0.12)	0.84	0.80	0.83	151.82
3 Factor	10077.06	74	<0.01	136.18	0.11 (0.11, 0.11)	0.86	0.83	0.86	136.18
6 Factor	8007.50	62	<0.01	129.15	0.11 (0.10, 0.11)	0.89	0.83	0.87	129.15
7 Factor	-	-	-	-	0.10 (0.10, 0.11)	0.90	0.84	0.90	123.08

Df = degrees of freedom; RMSEA = root mean square error of approximation; CFI = Comparative Fit Index; TLI = Tucker Lewis Index; NFI= Normed Fit Index

8.5 Discussion

The overall mean composite score based on the validation data was much lower than the study sample, which could be due to the fact that the study data was collected from the patients who were seeking dental treatment at the time of data collection while the reference data was collected from a representative sample of the general population. As the reference data had only very few missing values, the missing data were simulated using MCAR mechanism and also showed that the mean composite score did not differ significantly between missing and non-missing cases. Similarly, no significant correlation was observed between missingness and the demographic variables. The imputation of missing data for the reference data set using the methods discussed above showed similar results to the study sample and the mean values obtained using EM algorithm and multiple imputation methods are similar indicating the suitability of either of these methods to handle missing data in OHIP-14 items.

As in the study sample, the mean total score obtained using seven different missing data handling methods are similar.

Further analyses of missing data in the reference data set using Bland Altman method showed that the mean difference differed significantly for Regression and Multiple imputation methods compared to the original values and presence of fixed bias was observed in the data imputed using MI method. EM algorithm showed minimal difference with the original value as other methods (item mean, subject mean, interpolation and trend) with lower standard deviation and showed a random bias indicating the suitability of EM algorithm for handling missing data in OHIP-14.

A higher rate of floor effect and comparatively smaller ceiling effect were observed in the validation sample as in the study sample. This indicates that the majority of the individuals tend to give low scores while completing the OHIP-14 items. Both the ordinary least squares regression method and floor and ceiling effects adjusted Tobit model were fitted for the validation sample and showed that the magnitude of the effect of predictors (coefficient value) were larger in the Tobit model than in the OLS model. Similar results were observed in the study sample. This indicates that any analysis of OHIP data tends to underestimate the effect of predictors in the presence of floor and ceiling effects. The model comparison suggests that the Tobit model is the better fit, as previously shown in the study sample.

As in the study data, the construct validity of OHIP-14 was checked by fitting four different models (one, three, six and seven factor) using the reference data set. None of the four models showed good fit using the reference data set. The results were similar to the one obtained using study data set. The factors are highly correlated in the three, six and seven factor models indicating that the factors are not distinct. Some researchers have reported heavy loading on single component suggesting the existence of single factor in OHIP-14 items based on the factor analysis (Atchison and Dolan, 1990, Slade, 1997). A higher factor correlation in the three factor model has been reported (Santos et al., 2013). In this study the authors have concluded that OHIP-14 is one-dimensional and questioned the appropriateness of reporting results of OHIP-14 data as multi-dimensional. The correlation value exceeding 1 for a pair of factors in the seven factor model, lead to Heywood case. Therefore, the discriminant validity of these models is questionable.

The measurement of change could not be carried out with the validation data as this was a cross-sectional dataset. The results obtained using validation data for the other three methodological issues considered (missing data, floor and ceiling effects and factorial structure) are similar to the study data confirming the external validity.

9 General Discussion

9.1 Overview

OHIP-14 has been widely used in oral health research since its development in 1997. The reliability and validity of this instrument in various populations has been tested by many researchers. However there has been a lack of focus on the methodological problems that may arise while using this instrument to assess OHRQoL. Hence, the studies in this thesis have focussed on some of the methodological issues that may arise while using this instrument and the methods that may be used to address them in order to obtain accurate and precise estimates. . Though the original OHIP-14 questionnaire contained five responses (0="never", 1="hardly ever", 2="occasionally", 3="fairly often" and 4="very often") for each of the 14 items, six responses (0="never", 1="very little", 2="little", 3="much", 4="very much" and 5=intolerable") were adopted for each of the items in the present study. As the findings have been discussed under each section briefly in the previous chapters, this section provides an overview of all the findings, comparison with previous findings in the literature and the general discussion of their implications and suggestions for further studies.

Since this study addressed the methodological issues in OHIP-14, the reliability of this instrument in the study population was assessed by measuring the internal consistency of the items (the extent to which all of the items of an instrument measure the same latent variable) using Cronbach's alpha which was 0.94 showing high reliability and suitability of this instrument. This value for

OHIP-14 is similar to the one reported by Ikebe and co-workers (Ikebe et al., 2012) who have reported 0.95, Baba and co-workers (Baba et al., 2008) with 0.94 and Robinson and co-workers (Robinson et al., 2003) with 0.91. Locker and Slade (Locker and Slade, 1993) have suggested a value of 0.60 and above is considered to be good.

9.1.1 Missing data

The missing data mechanisms both in OHIP-14 responses and during follow-up were explored. The missing data in this study data were handled by using different missing data handling techniques and the results were compared. The multi-level modelling (Random effects model) was used to find out the significant predictors of OHIP total score. The focus on the methodological issues in using this instrument in measuring the OHRQoL in this study highlighted some valuable points for efficient use of this instrument. A missing rate of 23.06% in the OHIP-14 items was recorded in this study at baseline. This research has showed that the missing data in OHIP-14 items followed an MCAR pattern which facilitates researchers to justify the complete case analysis, provided the sample size is large enough. Though the item missingness in OHIP-14 follow the MCAR pattern in this study, use of complete case analysis reduces the power of the study as most of the research studies have sample size restrictions. There was a narrow range in the mean composite OHIP score based on the eight different missing data handling techniques tested, ranging from 21.46 to 22.81, indicating that all these techniques are similar in estimating the true value. Moreover, the sample size in this study was large enough to carry out complete cases analysis, as the data followed MCAR pattern. Though all the methods adopted here showed similar results, the Expectation Maximization (EM) algorithm and Multiple Imputation methods showed the closest approximations to the expected data. Moreover, as these two methods are considered to be statistically sound and use the maximum likelihood method of estimation which gives valid estimate, these two are the better options to handle this issue in case of smaller samples.

The comparison of original and imputed values for the missing data generated in the reference data set showed similar differences between the original and imputed values based on all the seven different methods. Further analyses using Bland and Altman graphs showed a random variability only for EM Algorithm, Regression, Subject mean and Multiple imputation methods. But the mean differences significantly differed for Regression and Multiple imputation methods. The significant mean difference and the line of equality not falling within 95% confidence interval of the mean difference indicate the possible existence of significant bias. The failure of line of equality falling within the confidence interval of the mean led to the conclusion that there is a significant systematic difference and hence these two methods either under- or over-estimate the true value (Giavarina, 2015). This study showed that the EM algorithm was a better choice among the seven methods to handle missing data in OHIP-14, as the difference between the imputed values using this method and the original values while analyzing the reference data set, are minimal with lower standard deviation and random bias.

As the present study involved three time points (baseline, 2 months and 4 months after treatment), the loss to follow up was inevitable. Therefore, the drop out mechanism was explored in the follow up data. Out of 360 completed baseline questionnaires, only 89 (24.72%) and 75 (20.83%) completed first and second follow up questionnaires. The OHIP scores did not influence the dropouts. However, the demographic variables namely age, gender and relationship did predict the drop outs in our study and hence there is less evidence for missing data to follow the MCAR pattern; it can be concluded that in the follow up, the missing data (loss to follow up) tend to follow the Missing At Random (MAR) pattern. Random effects models were used to handle this issue as this type of model can provide valid results even if the subject characteristics (age, gender and relationship) predicted the dropouts. The random effects models showed Education, Profession and Treatment needs as the significant predictors of OHIP scores. Though the temporal effect was prominent, it was not statistically significant. The observed decrease in total OHIP scores over the period of time cannot be fully attributed to treatment effect as this study does not have the control group for comparison.

The missing data in OHIP items ranged widely from 0.47% to 50% in different studies but no one has attempted to explore the missing data mechanism (Armfield and Ketting, 2015, Awad et al., 2014, Gisler et al., 2012, Kimura et al., 2012a, Reissmann et al., 2013). The methods used to handle missing data in OHIP-14 items have varied between researchers. While studying the Functional and Psychosocial impacts of oral disorders in Canadian adults, Locker and Quinonez substituted missing data with item mean and individuals with two or more missing values were deleted (Locker and Quiñonez, 2009). Similarly, Slade and co-workers have used sample means to handle the missing data (Slade et al., 2005). Many studies have either removed missing data from the analysis or substituted with sample mean/ median (Gagliardi et al., 2008, Hongxing et al., 2014, Hulme et al., 2016, Lahti et al., 2008). Though Motallebnejad and co-workers reported around 14% of missing data while validating the Persian version of OHIP-14, they adopted complete case analysis (Motallebnejad et al., 2011). The use of regression method to fill in the missing data was adopted by (Baba et al., 2008).

The multiple imputation was used by some researchers (Guarnizo-Herreño et al., 2014, Armfield and Ketting, 2015). Royston and co-workers (Royston, 2004) have recommended the MI method for handling missing data. The choice of missing data handling methods for the analysis of OHIP-14 data between EM algorithm and the MI method is mainly dependent on the analytical methods used and the availability of computation power. As multiple imputation requires more time and analytic power and also many statistical analyses are not available in the standard statistical packages for the data created from multiple imputation, the use of EM algorithm for handling missing data is highly recommended to achieve power. While analysing the missing data in using SF-36 used in Decennial health survey in France, Pyere and co-workers (Peyre et al., 2011) have confirmed multiple imputation and maximum likelihood as reference methods and suggested that these methods can be adopted to other questionnaires developed based on classical test theory. Similarly, Arnault and co-workers (Arnault et al., 2008) compared five imputation techniques namely

Monte-Carlo Markov Chain (MCMC), Expectation-Maximization (EM), Regression (REG), Propensity score (PROP) and Sequential regression (SEQ) and concluded that only Sequential regression method showed wider confidence interval and the changes varied depending on the imputation techniques. They have used Global Health Status Scale QLQ-C30, Utility Index EQ-5D and Visual analogue Scale (VAS) data. Though many investigators have used multiple imputation for handling missing data, it is recommended to explore the missing data mechanism first before selecting the method of imputing missing data.

The high dropouts in the follow-up while studying OHRQoL have been reported by many researchers. Saub and co-workers (Saub et al., 2005) have reported a response rate of 35.43% in the second data collection which was after two weeks interval and the response rate of 32.5% was achieved for the mail questionnaire after two week interval from the first data collection). Fueki and co-workers (Fueki et al., 2015), studying the effect of prosthetic treatment in the Japanese population, carried out at 3, 6 and 12 month follow ups after baseline using OHIP-49 and showed that 26.0% were lost to follow up/ post treatment evaluations. It is interesting to note that none of the studies have explored the dropout mechanisms in dental research. To assess the effect of treatment, various time points have been used by researchers. Awad and co-workers (Awad et al., 2014) have used baseline and 6 months after while studying the effect of mandibular 2 - implant over dentures on OHRQoL. Oliveira and co-workers (Douglas de Oliveira et al., 2013) have found a significant decline in OHIP-14 scores after 3 months of treatment in Physical pain and disability and psychological discomfort and disability dimensions. Reissman and co-workers (Reissmann et al., 2013) reported an increase in the total OHIP scores from baseline to first follow up (after 3 days of surgery) and a decrease in the second follow up (after 4 weeks). While studying the suitability of OHIP-14 for assessing the OHRQoL and its responsiveness, Baba and co-workers (Baba et al., 2008) used a recall period of one month and found that the average change score from baseline to follow up was 5.2. A study carried out by Enoki and co-workers (Enoki et al., 2013) showed that there was no significant improvement in OHRQoL of Japanese elderly population after 7 years of follow up. The need for longer follow up periods particularly when longer treatments are required while measuring the OHRQoL of dental patients have been stressed by Fueki and co-workers (Fueki et al., 2015). Similarly, lengthy follow up periods have been suggested by Lodi and co-workers (Lodi et al., 2012) in order to evaluate the long term benefit of interventions using OHRQoL measures.

9.1.2 Floor and Ceiling Effects

The presence of floor and ceiling effects was confirmed and quantified using the method suggested by Van Den Oord and Van Der Ark (Oord and Ark, 1997). The findings from the OLS model was then compared with Tobit model – a model in which the floor and ceiling effects are adjusted. The floor effect which can be termed as Average Item Difficulty (also termed as item sufferings) was 0.4 indicating that 40% of subjects consistently tend to give 0 for the items due to the floor effect of OHIP items and have a censored scale. Similarly, the ceiling effect, which is termed as Average item easiness was 0.03, indicating that only 3% of the subjects consistently tended to give the maximum score for the items due to the ceiling effect of OHIP items. This showed that items in OHIP-14 suffer from both floor and ceiling effects. However, the floor effect is very prominent and the ceiling effect is minimal.

The floor and ceiling values for OHIP-14 composite score were estimated to be 14 and 56 respectively. The OLS model identified that students and retired people have significantly lower scores compared to unemployed people whereas Tobit model after adjusting for floor and ceiling effects identified that only retired people had a significantly lower score when compared to unemployed. The model that was adjusted for floor and ceiling effects showed an increase in the magnitude of the predictor coefficients indicating that the model underestimated the effect of predictors on OHIP scores when floor and ceiling effects were not taken into consideration. Particularly when the sample size is small, these effects tend to change the overall results of the analysis and hence it is important to consider the floor and ceiling effects while analysing OHIP data.

As in the present study, prominent floor effects and minimal ceiling effects have been reported by other workers (Locker et al., 2001, Stenman et al., 2012). Locker and co-workers stated that around 30% of the respondents scored 0 in the additive method and 45.8% in the count method. Ohrn and co-workers (Öhrn and Jönsson, 2012), while comparing GOHAI and OHIP-14 for measuring OHRQoL of periodontal patients before and after dental hygiene treatment, found that the floor effect was more prominent in OHIP-14 than GOHAI. Apart from studies of OHRQoL, the presence of floor and ceiling effects have been found by many researchers in other areas of study. Flolo and co-workers (FløIo et al., 2014) while evaluating the Norwegian version of an American short form version (WEL-SF) of the commonly used Weight Efficacy Lifestyle Questionnaire (WEL), reported an average of 0.9 and 17% of ceiling effects in non-operated and operated samples respectively. Less than 12% floor and ceiling effects were reported by Teixeria and co-workers (Teixeira et al., 2012) while evaluating the Portuguese version of the Kiddo-KINDL questionnaire in Brazil to measure the QoL of children. Allen and Locker have reported the potential for floor effect when using the current OHIP-14 and found that this affected the measurement properties when used for assessment of edentulous patients (Allen and Locker, 2002). Maindal and co-workers (Maindal et al., 2012) reported a high range of both floor and ceiling effects (floor effect: 2.9 to 69.2%; ceiling effect 4.0 to 40.4%) in the Danish version of the Patient Assessment of Chronic Illness Care (PACIC) 20-item questionnaire which is used to measure how chronic care patients perceive their involvement in care. Similarly, high percentage of both floor and ceiling effects were observed in two dimensions by Yao and co-workers (Yao et al., 2010) in China, while evaluating the reliability and validity of SF-36 in patients with advanced schistosomiasis. In AIDs research, Soarez and co-workers (Soárez et al., 2009) while evaluating the Brazilian version of the HIV/AIDS-Targeted Quality of Life Instrument (HAT-QoL) observed the ceiling effect in 7 out of 9 dimensions with the highest value of 63.2%. The same study has reported a floor effect of 30.2% for financial worries. The presence of floor effects and a large percentage of ceiling effects were also reported in the quality of the functional scales used to assess patients with different types of Muscular Dystrophy MD (Lue et al., 2009). Jenkinson and co-workers (Jenkinson et al., 1996) while measuring the sensitivity of short form of SF-36 health status measure, reported a lower floor effect but higher percentage of ceiling effect in their study. Though many investigators have reported the presence of floor / ceiling effects in their data, none have attempted to quantify and adjust these effects in their analyses.

Therefore, it is recommended from the current analyses that studies which use OHIP-14 for measuring the OHRQoL should address the floor and ceiling effects using the Tobit model instead of OLS regression.

9.1.3 Change scores

The measurement of treatment effect involves follow up of patients over a period of time after the intervention. As the patients are followed over a period of time, this type of study encounters many issues both at data collection (loss to follow up) and at the analysis stage (handling missing cases and responsiveness). While the other two (loss to follow up and handling missing cases) have already been discussed above, we also investigated the responsiveness of OHIP-14 for change. This has the practical importance as it helps to decide on the treatment strategy and future planning. The Reliability of Change (RC) measured in this study was higher for baseline to first and second follow ups. Similarly, the reliability of change values based on the Reliability of Change Index (RCI), an improved method for measuring the reliable and clinically significant change, were 0.83 and 0.56 for baseline to second and first to second follow ups respectively. Based on these figures, individuals were classified into any one of three groups namely, reliable decline (Worsened), no change and reliable improvement (Improved). Participants were classified into one of these three groups both by using RCI and Standard Error of Measurement. From baseline to second follow up, 13.33% of participants showed an improvement using RCI while it was 30.67% when we used SEoM. It is interesting to note that 4.00% have worsened using RCI and 29.33% have worsened using SEoM, as most of the participants in this study are referral patients who need long term treatment, it could be argued that the study period of 4 months after initial treatment could not be sufficient to measure the full effect of treatment. As discussed above, longer study periods are recommended to study the full treatment effect particularly when the long term treatments are needed.

In the process of exploring responsiveness of OHIP-14 for change, many investigators have attempted to measure the change and the method adopted to measure the change from baseline to follow-up differed between studies. The traditional methods used are the comparison of mean scores and the use of effect size. However, this gives only the overall change and it is important to measure the individual change as the treatment for oral problems differs between individuals. The effect size used to measure the change in OHRQoL varied between investigators. Yule and co-workers (Yule et al., 2015) have used effect size to measure the responsiveness to change and reported that it was 0.4 for both OHIP-49 and shortened version. The mean change score has been used to assess the minimally important clinical difference which is a crude method of measuring it. The use of OHIP-49 has been criticised for the presence of redundant items which may affect the responsiveness to change (Durham et al., 2011, John et al., 2002, Larsson et al., 2004, Segù et al., 2005). The commonly used methods for measuring the Minimal Important Difference (MID) are Effect size and Standard Error of Measurement (SEoM) (Durham et al., 2011, John et al., 2002, Larsson et al., 2004, Revicki et al., 2008, Segù et al., 2005, Tsakos et al., 2012, Wyrwich et al., 2013). Using repeated measures ANOVA to measure the responsiveness of OHIP-14 while measuring the OHRQoL for pre- and post-operative time points, it was concluded that OHIP-14 is internally responsive to changes in impacts of oral conditions (Kieffer et al., 2012). In this study, the investigators have used partial eta square as a measure of effect size and have reported a large effect size (partial eta square value of 0.67).

Fueki and co-workers (Fueki et al., 2015) reported a moderate decrease in OHIP-14 composite score from baseline to 12 months follow up and observed a change of 8.2 in the composite score from baseline to follow up which is higher than the minimally important difference suggested by John and co-workers (John et al., 2009) for prosthetic treatment. The present study observed an effect size of 0.2 from baseline to second follow up which is considered to be low. A similar low effect size of 0.2 in a conventional denture group and 0.3 in an implant group using OHIP-14 has been reported in a study for developing a shorter version of Oral Health Impact Profile in assessing oral health related quality of life in edentulous patients using Canadian and UK data (Allen and

Locker, 2002). Stone and co-workers (Stone et al., 2015) while studying a plaque control intervention in a UK population, recorded a medium effect size from baseline to two weeks and twenty weeks follow ups. Locker (Locker et al., 2004) and co-workers while studying the oral health of elderly patients established a change of 5 points whereas Allan and co-workers (Allen et al., 2009) showed a change of between 7 to 10 points for the minimal important difference. Jonsson and co-workers (Jönsson and Öhrn, 2014) found a change of 3.9 points in the OHRQoL-UK (16 item questionnaire) scores with three months reference period and observed an effect size of 0.3 with SEoM 2.3 in a Swedish sample. None of the studies have attempted to measure the individual change by using different methods. The change in the HRQoL of eating disorder patients has been discussed by Padierna and co-workers (Padierna et al., 2002). The reliability of differences between pre- and post- test scores has been proved by Zimmerman and Williams (Zimmerman and Williams, 1982). Clinically meaningful change has been widely discussed in Mental Health treatment by Eisen and co-workers (Eisen et al., 2007).

There are varying opinions about the follow-up period required while measuring the change. The need for a longer follow-up period (minimum of 6 months) has been suggested by Gupta (Gupta, 2004). The ability of OHIP-14 to measure the complete change is questionable in this study as a lower effect size is recorded which is similar to the findings of Allen and Locker (Allen and Locker, 2002) who reported that responsiveness to change was not good for OHIP-14. Allen and co-workers (Allen et al., 2009) have stressed the need to have a bench mark to interpret the analysis of results. They have concluded that a change in total score of between 7 and 10 while using OHIP-20 are considered as the minimal important difference. The method suggested here is an improved method and facilitates to classify the patient into one of three categories discussed. The ability to classify the patients into one of the three groups helps to measure the success of the treatment given.

9.1.4 Factorial structure

The attempt to explore the existence of the number of dimensions in the OHIP-14 items revealed some interesting findings which should be considered while analysing the OHIP data. The comparison of four different models considered in this study using the study data revealed that none of the models fitted the data well. As OHIP-14 was developed based on Locker's theoretical model it is important that the OHIP-14 items fall into the domains explained in Locker's model. The comparison of four models was based on confirmatory factor analyses using Structural Equation modelling. Though the original developer of OHIP-14 identified seven factors using factor analysis, they also reported high correlation between the factors (Slade, 1997). The high correlation between the factors suggests that the OHIP total score may be treated as a single score for further analyses instead of dealing with seven factors separately.

In the comparison of four different models that have been used in OHIP-14 analyses, one of the correlations between the factors showed more than the maximum value of 1 for the seven factor model, indicating that this model encountered a Heywood case (Rindskopf, 1984). The causes of a Heywood case have been widely discussed and may be due to many reasons. The presence of outliers in the data which leads to a Heywood case has been reported by Bollen (Bollen, 1987). The problem of non-convergence and under-identification of the model has been reported by Van Driel and Boomsma and Hoogland (Van Driel, 1978, Boomsma and Hoogland, 2001). Rindskopf (Rindskopf, 1984) reported empirical under-identification while others (Bollen, 1989, Dillon et al., 1987, Kolenikov and Bollen, 2012, Sato, 1987, Van Driel, 1978) attributed structurally mis-specified models for Heywood case. Some researchers have reported sampling fluctuations as one of the reasons that lead to Heywood case (Anderson and Gerbing, 1984, Boomsma, 1983, Van Driel, 1978).

The identification of the presence of a Heywood case in the seven factor model indicated that either this model is inappropriate for OHIP-14 or the model is structurally mis-specified. Such situations must be carefully looked into as this may arise either because the model does not support the theoretical framework based on which the instrument was developed, or the number of items in one or more factors may not sufficiently represent the unknown construct which we intend to measure. As OHIP-14 contained only 14 items and the theoretical model defined seven dimensions with two items per factor, it could be argued that either the number of items per factor may not be sufficient or the number of factors could be less than 7 which might have led to Heywood's case. This finding is supported by that obtained by Baker and co-workers (Baker et al., 2008b) who used OHIP-49 and arrived at six factors based on the revised 22 items, which was a better fit to the data. The high correlation between the factors and the analysis leading to a Heywood case confirms that either the existence of seven factors in OHIP-14 is questionable or the items included in the instrument are not exhaustive and other important items are missing. Therefore, the current analyses show considerable evidence for the unsuitability of seven factor model while handling OHIP-14 data. As Gregory and co-workers (Gregory et al., 2005) have arrived with seven factors based on the qualitative research which included oral health behaviour along with oral health knowledge, further research is required by including oral health behaviour in the assessment of OHRQoL.

The current analyses also found that all the four models tested did not fit both the study data and the reference data set. Though some investigators (John et al., 2014, Santos et al., 2013) have reported that OHIP-14 items should be analysed as a single factor, the current analyses do not support this. Similarly, the three, six and seven factor models failed to fit both the data set. The Baker and co-workers' (Baker et al., 2008b) six factor model (based on revised 22 items from OHIP-49) included some items from OHIP-49 which are not in OHIP-14 which could be the reason for this model not fitting this data set. This indicates that the current items in OHIP-14 do not represent adequately Locker's theoretical model. Therefore, this study suggests the necessity for further work on the theoretical aspects of OHRQoL as suggested by Baker and co-workers.

9.1.5 Comparison of the findings from the current study and National sample

The methodological issues discussed in the research and the methods suggested from our results were subsequently tested using an independent dataset collected nationally for the Adult Dental Health Survey, as a validation study which aimed to give our findings increased external validity. As this study does not have missing data, missing data were generated randomly for comparison. Since both the actual data and the data generated by various missing data techniques were available in this data set, it is an added advantage to make comparisons with the actual data. The analyses of this validation sample produced similar results in line with the findings of studies on the primary cohort, indicating the reliability and validity of these methods to a wider and more heterogeneous population. As the validation data were collected from the general population while the study samples were taken from patients at the time of treatment, a high percentage of left censoring was observed in the validation data. The average item difficulty was higher (88.00%) when compared to the original sample (40.00%) which could be due to the fact that the data were collected from the general population whereas the study data were collected from patients who came for treatment at the time of survey. As in the original analysis, the floor and ceiling effects-adjusted Tobit model showed a greater magnitude for all the predictors when compared to the traditional OLS model, indicating the value for adjusting floor and ceiling effects in OHIP research. Though none of the predictors were significantly different for the overall composite score between Tobit and OLS models, the dimension wise analysis showed some differences between these two models. The responsiveness of OHIP-14 could not be carried out in the validation data set as it is a cross sectional survey and hence no follow up were carried out. The testing of four models namely one, three, six and seven factor models produced similar results as in the study data. None of the models fitted the validation data and ended with Heywood case. The comparison of the findings of the present study with the results obtained by analysing national data set revealed similar results indicating the reliability and validity of the present findings.

9.2 Strengths and Weaknesses of this study

The main strength of any study is its data. The use of a primary data may have several advantages over secondary data, particularly that the primary data is collected with a specific aim and the power and effect size are considered in the sample size calculation for the specific objectives. As the primary aim of this study was to address the methodological issues, the present study data is an added strength to this study. As this study used primary data collection, the study was well planned so that the research questions can be addressed with accuracy. The validation of the findings of a study with another data set provides external validity and confirms the methods and the results. As the methods employed to address the four methodological issues and the results in this study are validated using an independent dataset with large sample size collected for the England, Wales and Northern Ireland Adult Dental Health Survey, this provides external validity for the findings. Also, as this study involved missing data and the validation sample had no missing data, the comparison of these two studies while addressing techniques used to handle missing data is more appropriate in comparing the parameters based on estimated and original values.

The collected data contained a relatively higher percentage of missing data which is the main weakness of this study as far as the secondary aim of measuring OHRQoL is concerned. Another important limitation of this study is that the samples included only the patient groups and did not have a control group. Hence, the effect observed in this study could not be attributed to treatment as there is no control group. The study samples contained people suffering from various dental conditions and the dental treatments are markedly different for each type of condition. Previous research has also shown that the OHRQoL differs according to different dental problems (Cimilli et al., 2012) . So, it might have been better if the diseases were analysed separately depending on the hypotheses being tested. On the other hand, in the present study the combined data was used for the analysis, which does give a broader perspective to the effect of dental conditions that is not restricted to specific conditions. Moreover, as the treatment options were described using dental terms in the questionnaire such as orthodontics, endodontics etc. many of the patients might not have understood this and hence many people did not answer this item. Therefore, further disease specific analyses could not be carried out.

Another important issue is that the data were collected from patients who attended King's College Hospital where the majority were patients referred by general dentists to the hospital for treatment. Usually referral patients will tend to have some complicated problems and may need long term treatment. However, the change scores are measured only by considering the maximum follow up period of four months which is a relatively shorter period of time when compared to the severity of the illness and the treatment duration. Moreover, only two follow-ups were taken in this study but the long term effect of treatment can be measured only if more and longer follow-ups are considered (Lodi et al., 2012). Therefore, the full effect of treatment could not be measured. The catchment area is limited in this study as the samples are taken from King's College Hospital only where the majority of the patients were living around London area which is not representative of the entire UK population. Because the sample taken in this study is more homogeneous, the findings of this study cannot be extended to the entire UK population. Also, the suitability of the analytical methods suggested in this study to the heterogeneous population has to be checked. It is also important to cover large areas containing heterogeneous samples in terms of economic condition, ethnicity, educational level, treatment availability etc. Also, this study has not taken clinical aspects in to account and hence the comparison of clinical variables with the OHRQoL could not be possible.

9.3 Future work

Further work in this area would be valuable both for dealing with methodological issues that arise when assessing OHRQoL of dental patients in addition to further studies on the effect of dental attendance on OHRQoL. As the primary aim of this work was to address the methodological issues while measuring

OHRQoL using OHIP-14, the methods suggested to address the issues should be checked further their application in other studies. The influence of measurement scales for each item of an instrument on the floor and ceiling effects is a key issue when considering an instrument for measuring an unknown construct. So, it would be worth to consider a study (using OHIP-14) with more than one type of scale (for example Likert scale and VAS score ranging from 0 to 100) and compare the floor and ceiling effects, so that the influence of type of scale while measuring OHRQoL could be assessed. The present study has encountered Heywood's case while evaluating the number of dimensions in the OHIP items which questions either the validity of the number of items in each factor or the number of factors in OHIP-14 itself. Moreover, the study by Gregory and co-workers have suggested seven dimensions based on relevance framework and Baker and co-workers have stated the presence of six dimensions while measuring OHRQoL using OHIP which include items in original OHIP apart from OHIP-14. Therefore, future work incorporating all the 49 items in the original OHIP along with the items related to Oral Health behavioural aspects in OHRQoL measure may be required to validate the theoretical framework and to evaluate the number of dimensions.

Although it was not a primary aim of this study to assess the effect of dental attendance on OHRQoL, the results obtained do partly provide evidence of this issue. As the present study was mainly based on patients attending King's College Hospital for dental treatment it would not be appropriate to extend this to a more general population and hence future studies considering a heterogeneous sample representing the entire UK population based on suitable sampling techniques are required to check the suitability of the findings to address both the methodological issues and issues of impacts of dental interventions. As the OHRQoL differs according to dental conditions, particularly with aesthetic aspects being involved, future work is needed to check these methodological issues and to assess the OHRQoL of dental patients for different dental conditions by using a reasonably large sample size so that each dental condition will have enough samples for the analysis. In future studies, samples could be selected from a range of clinics covering wider catchment areas or from primary dental care settings with heterogeneous populations representing the whole country. This study has not taken consideration of the clinical conditions or the clinical aspects in the analysis. As the OHRQoL differs according to clinical conditions and the type of treatments received, studies incorporating the clinical aspects are required to measure the OHRQoL. Also this study has considered only three time points with two months interval to measure the change in OHRQoL over the period of time. But most of the samples considered in this study are referral patients with long term treatment, future work involving more time points to address the long-term effect of treatment and the residual effect are needed.

9.4 Summary and Conclusions

The present research has shown that the missing data in OHIP-14 items tend to follow MCAR pattern and hence complete case analysis may be used. However, considering the sample size, EM algorithm is recommended to handle missing data based on the current research. There is a high chance of missing data (drop outs) in the follow up to follow missing at random pattern and therefore further analyses may be carried out using statistical techniques which handle this issue. As multi-level models (random effects models) take all the data into account and address this issue, it is suggested to use this technique for further analyses of follow up data when the data contains considerable amount of missing data (usually more than 5%). This study confirmed that OHIP-14 data suffer from floor and ceiling effects and the floor effect is more prominent. These two effects are quantified by using interpolation and this method may be adopted to quantify these effects in future studies. The floor and ceiling effects are adjusted using tobit regression and it is also proved that this model gives improved estimates for the parameters. As the presence of floor and ceiling effects in OHIP data is confirmed, any future study must take these into account in order to get improved estimates for the unknown parameters. Though varying opinions prevail among researchers regarding the number of factors in OHIP-14 items, this study has shown that none of the four models considered in this research fitted this data. This indicates the necessity to look into the theoretical aspects of OHIP-14 in measuring the Oral Health Related Quality of Life. The occurrence of Haywood's case in the seven factor model, raises the concern over the use of this model in analysing OHIP-14 data to measure the OHRQoL. This study has shown a better way of quantifying the change in oral health research rather than using the conventional measure of effect size. As the measure of change is important in treatment planning and to measure the effect of treatment, this study has given an improved measure for this. This facilitates the researcher to classify the patients into one of the three categories (Worsened, No Change and Improved). In this study, the four methodological issues taken are explored using the study data and the methods to handle these issues while using OHIP-14 are discussed. The findings of this study are validated using the national data base which confirms the external validity of these methods.

In conclusion, OHRQoL measured using OHIP-14 instrument does have some methodological issues which are to be considered when analysing data in order to improve the power of the study and the accuracy of the results. The missing items in OHIP-14 tend to follow MCAR pattern and hence the missing data may be ignored from the analysis if the sample size is large or imputed using EM algorithm. As this method provides better estimates for the missing data, it is suggested to use EM algorithm as this is comparatively easier and all standard statistical packages can handle all the relevant statistical analysis. As the dropouts in the follow up studies tend to follow MAR pattern, it is strongly suggested to used Random effects models as these types of models use all the available data and provide valid results even under less restrictive assumption of MAR. OHIP-14 data do suffer from floor and ceiling effects which have a significant impact on the predictors and hence these effects must be adjusted. The suggested method in this research helps to quantify these effects and adjust them in the model using Tobit regression instead of the traditional regression model which underestimates the effect.

As measuring the effectiveness of dental treatments is important, it is essential to measure the true effect using improved methods. This research has used a relatively new index namely Reliable Change Index which helps to measure the actual overall change and also define a bench mark to categorise an individual under one of the three groups namely ‘Improved’, ‘No change’ and ‘Worsened’. This will really help the researchers to plan the future treatment strategy for individuals. Though the number of factors in OHIP-14 items may vary according to studies, none of the models considered in this study (one, three, six and seven factor models) fitted this data and the seven factor model recommended by the original developer of OHIP-14 ended with Heywood case. This indicates either the existence of seven factors is questionable or the number of items in each factor is not appropriate. Moreover, the factors are highly correlated with one another suggesting that any analysis based on the factors should be carefully handled and interpreted. This not only provides suitable analytical methods for methodological issues considered in this field but also to measure the treatment effect more efficiently by addressing these issues. Hence, based on the analyses of the study and the reference data set, the aims and the nine objectives of this research have been met.

REFERENCES

AINAMO, J., BARMES, D., BEAGRIE, G., CUTRESS, T., MARTIN, J. & SARDO-INFIRRI, J. 1982. Development of the World Health Organization (WHO) community periodontal index of treatment needs (CPITN). *International Dental Journal,* 32**,** 281-291.

AL SHAMRANY, M. 2006. Oral health-related quality of life: a broader perspective. *Eastern Mediterranean Health Journal,* 12**,** 895.

AL SUDANI, F. Y., VEHKALAHTI, M. M. & SUOMINEN, A. L. 2016. Association of current employment status with oral health related behaviors: findings from the Finnish Health 2000 Survey. *European Journal of Oral Sciences,* 124**,** 368-376.

ALLEN, F. & LOCKER, D. 2002. A modified short version of the oral health impact profile for assessing health-related quality of life in edentulous adults. *International Journal of Prosthodontics,* 15.

ALLEN, P., THOMASON, J., JEPSON, N., NOHL, F., SMITH, D. & ELLIS, J. 2006. A randomized controlled trial of implant-retained mandibular overdentures. *Journal of Dental Research,* 85**,** 547-551.

ALLEN, P. F. 2003. Assessment of Oral Health Related Quality of Life. *Health and Quality of Life Outcomes,* 1**,** 40.

ALLEN, P. F., MCMILLAN, A. S. & LOCKER, D. 2001. An assessment of sensitivity to change of the Oral Health Impact Profile in a clinical trial. *Community Dentistry and Oral Epidemiology,* 29**,** 175-182.

ALLEN, P. F., O'SULLIVAN, M. & LOCKER, D. 2009. Determining the minimally important difference for the Oral Health Impact Profile-20. *European Journal of Oral Sciences,* 117**,** 129-134.

AMERICAN DENTAL ASSOCIATION (ADA) 2014. ADA Policy - Definition of Oral Health. [Available from: http://www.ada.org/en/about-the-ada/ada-positions-policies-and-statements/ada-policy-definition-of-oral-health. Accessed on: 10th May 2017].

ANDERSON, J. C. & GERBING, D. W. 1984. The effect of sampling error on convergence, improper solutions, and goodness-of-fit indices for maximum likelihood confirmatory factor analysis. *Psychometrika,* 49**,** 155-173.

ANDERSON, L. A. & DEDRICK, R. F. 1990. Development of the trust in physician scale: A measure to assess interpersonal trust inpatient-physician relationships. *Psychological Reports,* 67**,** 1091-1100.

ARCHIE JR, J. P. 1981. Mathematic coupling of data: a common source of error. *Annals of Surgery,* 193**,** 296.

ARMFIELD, J. M. & KETTING, M. 2015. Predictors of dental avoidance among Australian adults with different levels of dental anxiety. *Health Psychology,* 34**,** 929.

ARNAULT, A., IVANESCU, C., VAN ENGEN, A. & PEETERS, P. 2008. PCN109 Dealing with Quality of Life missing data in a single arm study. Comparison of Multiple Imputation methods *Value in Health,* 11**,** A494-A495.

ARRINDELL, W. A. & VAN DER ENDE, J. 1985. An empirical test of the utility of the observations-to-variables ratio in factor and components analysis. *Applied Psychological Measurement,* 9**,** 165-178.

ÅSTRØM, A., HAUGEJORDEN, O., SKARET, E., TROVIK, T. & KLOCK, K. 2005. Oral Impacts on Daily Performance in Norwegian adults: validity, reliability and prevalence estimates. *European Journal of Oral Sciences,* 113**,** 289-296.

ÅSTRØM, A., HAUGEJORDEN, O., SKARET, E., TROVIK, T. & KLOCK, K. 2006. Oral Impacts on Daily Performance in Norwegian adults: the influence of age, number of missing teeth, and socio-demographic factors. *European Journal of Oral Sciences,* 114**,** 115-121.

ÅSTRØM, A. N., EKBÄCK, G. & ORDELL, S. 2010. Factor structure of a conceptual model of oral health tested among 65 year olds in Norway and Sweden. *Community Dentistry and Oral Epidemiology,* 38**,** 110-119.

ATCHISON, K. A. & DOLAN, T. A. 1990. Development of the Geriatric Oral Health Assessment Index. *Journal of Dental Education,* 54**,** 680-687.

AWAD, M., LOCKER, D., KORNER-BITENSKY, N. & FEINE, J. 2000. Measuring the effect of intra-oral implant rehabilitation on health-related quality of life in a randomized controlled clinical trial. *Journal of Dental Research,* 79**,** 1659-1663.

AWAD, M. A., LUND, J. P., SHAPIRO, S. H., LOCKER, D., KLEMETTI, E., CHEHADE, A., SAVARD, A. & FEINE, J. S. 2003. Oral health status and treatment satisfaction with mandibular implant overdentures and conventional dentures: a randomized clinical trial in a senior population. *International Journal of Prosthodontics,* 16.

AWAD, M. A., RASHID, F. & FEINE, J. S. 2014. The effect of mandibular 2-implant overdentures on oral health–related quality of life: an international multicentre study. *Clinical Oral Implants Research,* 25**,** 46-51.

BABA, K., INUKAI, M. & JOHN, M. 2008. Feasibility of oral health-related quality of life assessment in prosthodontic patients using abbreviated Oral Health Impact Profile questionnaires. *Journal of Oral Rehabilitation,* 35**,** 224-228.

BAE, K. H., KIM, H. D., JUNG, S. H., PARK, D. Y., KIM, J. B., PAIK, D. I. & CHUNG, S. C. 2007. Validation of the Korean version of the oral health impact profile among the Korean elderly. *Community Dentistry and Oral Epidemiology,* 35**,** 73-79.

BAKER, S., GIBSON, B. & LOCKER, D. 2008a. Is the oral health impact profile measuring up? Investigating the scale's construct validity using structural equation modelling. *Community Dentistry and Oral Epidemiology,* 36**,** 532-541.

BAKER, S. R., PANKHURST, C. L. & ROBINSON, P. G. 2006. Utility of two oral health-related quality-of-life measures in patients with xerostomia. *Community Dentistry and Oral Epidemiology,* 34**,** 351-362.

BAKER, S. R., PEARSON, N. K. & ROBINSON, P. G. 2008b. Testing the applicability of a conceptual model of oral health in housebound edentulous older people. *Community Dentistry and Oral Epidemiology,* 36**,** 237-248.

BARBOSA, T. & GAVIÃO, M. 2008. Oral health-related quality of life in children: Part II. Effects of clinical oral health status. A systematic review. *International Journal of Dental Hygiene,* 6**,** 100-107.

BENNADI, D. & REDDY, C. 2013. Oral health related quality of life. *Journal of International Society of Preventive & Community Dentistry,* 3**,** 1.

BENNETT, D. A. 2001. How can I deal with missing data in my study? *Australian and New Zealand Journal of Public Health,* 25**,** 464-469.

BENNETT, M. & PHILLIPS, C. 1998. Assessment of health-related quality of life for patients with severe skeletal disharmony: a review of the issues. *The International Journal of Adult Orthodontics and Orthognathic Surgery,* 14**,** 65-75.

BENSON, P. E., CUNNINGHAM, S. J., SHAH, N., GILCHRIST, F., BAKER, S. R., HODGES, S. J. & MARSHMAN, Z. 2016. Development of the Malocclusion Impact Questionnaire (MIQ) to measure the oral health-related quality of life of young people with malocclusion: part 2–cross-sectional validation. *Journal of Orthodontics,* 43**,** 14-23.

BENSON, P. E., DA'AS, T., JOHAL, A., MANDALL, N. A., WILLIAMS, A. C., BAKER, S. R. & MARSHMAN, Z. 2014. Relationships between dental appearance, self-esteem, socio-economic status, and oral health-related quality of life in UK schoolchildren: A 3-year cohort study. *The European Journal of Orthodontics***,** cju076.

BOLLEN, K. A. 1987. Outliers and Improper Solutions A Confirmatory Factor Analysis Example. *Sociological Methods & Research,* 15**,** 375-384.

BOLLEN, K. A. 1989. Structural equations with latent variables. *John Wiley & Sons.*

BOOMSMA, A. 1983. On the robustness of LISREL (maximum likelihood estimation) against small sample size and nonnormality. *Groningen: s.n. Doctoral Thesis*.

BOOMSMA, A. & HOOGLAND, J. J. 2001. The robustness of LISREL modeling revisited. *Structural Equation Models: Present and Future. A Festschrift in honor of Karl Jöreskog***,** 139-168.

BORTHWICK-DUFFY, S. A. 1992. Quality of life and quality of care in mental retardation. *Mental Retardation in the Year 2000.* Springer.

BOWLING, A. 2005. Just one question: If one question works, why ask several? : BMJ Publishing Group Ltd.

BRENNAN, D. S. & SPENCER, A. J. 2004. Dimensions of oral health related quality of life measured by EQ-5D+ and OHIP-14. *Health and Quality of Life Outcomes,* 2**,** 1.

BRIGGS, A., CLARK, T., WOLSTENHOLME, J. & CLARKE, P. 2003. Missing.... presumed at random: cost-analysis of incomplete data. *Health Economics,* 12**,** 377-392.

BRODER, H. L., SLADE, G., CAINE, R. & REISINE, S. 2000. Perceived impact of oral health conditions among minority adolescents. *Journal of Public Health Dentistry,* 60**,** 189-192.

BROWN, R. 1989. Congeneric modeling of reliability using censored variables. *Applied Psychological Measurement,* 13**,** 151-159.

BROWN, T. 2006. Confirmatory factor analysis for applied research. *Second edition. Guilford Press; New York*.

BRUNNER, E. J. & MARMOT, M. G. 2005. Social organisation, stress and health. *In: Marmot, MG and Wilkinson, R and G, , (eds.) Social Determinants of Health. (pp. 17-43). Oxford University Press: Oxford. .*

CALMAN, K. C. 1984. Quality of life in cancer patients--an hypothesis. *Journal of Medical Ethics,* 10**,** 124-127.

CARR, A. J., GIBSON, B. & ROBINSON, P. G. 2001. Is quality of life determined by expectations or experience? *British Medical Journal,* 322**,** 1240.

CATELL, R. 1978. The scientific use of factor analysis. *Behavioural and Life Sciences. Plenum Press. New York and London***,** 205-238.

CHEUNG, K., OEMAR, M., OPPE, M. & RABIN, R. 2009. EQ-5D user guide: basic information on how to use EQ-5D. *Rotterdam: EuroQol Group*.

CIMILLI, H., KARACAYLI, U., ŞIŞMAN, N., KARTAL, N. & MUMCU, G. 2012. Comparison of the oral health-related quality of life and dental pain in symptomatic irreversible pulpitis and pericoronitis. *Journal of Dental Sciences,* 7**,** 250-260.

COLLINS, L. M., SCHAFER, J. L. & KAM, C.-M. 2001. A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods,* 6**,** 330.

COMREY, A. & LEE, H. 1992. A First Course in Factor Analysis. *Psychology Press; 2 edition (January 22, 2016).*

CORNELL, J. E., SAUNDERS, M. J., PAUNOVICH, E. D. & FRISCH, M. B. 1997. Oral health quality of life inventory (OH-QoL). *Measuring Oral Health and Quality of Life***,** 136-149.

CORRIDORE, D., CAMPUS, G., GUERRA, F., RIPARI, F., SALE, S. & OTTOLENGHI, L. 2013. Validation of the Italian version of the Oral Health Impact Profile-14 (IOHIP-14). *Annali di Stomatologia,* 4**,** 239.

CROCOMBE, L. A., BRENNAN, D. & SLADE, G. 2012. The influence of dental attendance on change in oral health–related quality of life. *Community Dentistry and Oral Epidemiology,* 40**,** 53-61.

CUNNINGHAM, S. J., GARRATT, A. M. & HUNT, N. P. 2000. Development of a condition-specific quality of life measure for patients with dentofacial deformity: I. Reliability of the instrument. *Community Dentistry and Oral Epidemiology,* 28**,** 195-201.

CUNNINGHAM, W. A., PREACHER, K. J. & BANAJI, M. R. 2001. Implicit attitude measures: Consistency, stability, and convergent validity. *Psychological Science,* 12**,** 163-170.

DAHL, K. E., WANG, N. J., SKAU, I. & ÖHRN, K. 2011. Oral health-related quality of life and associated factors in Norwegian adults. *Acta Odontologica Scandinavica,* 69**,** 208-214.

DE GOEIJ, M. C., VAN DIEPEN, M., JAGER, K. J., TRIPEPI, G., ZOCCALI, C. & DEKKER, F. W. 2013. Multiple imputation: dealing with missing data. *Nephrology Dialysis Transplantation,* 28**,** 2415-2420.

DE OLIVEIRA, C. & SHEIHAM, A. 2004. Orthodontic treatment and its impact on oral health - related quality of life in Brazilian adolescents. *Journal of Orthodontics,* 31**,** 20-27.

DEYO, R. A., DIEHR, P. & PATRICK, D. L. 1991. Reproducibility and responsiveness of health status measures statistics and strategies for evaluation. *Controlled Clinical Trials,* 12**,** S142-S158.

DILLMAN, D. 1978. Mail and telephone surveys: The total design method. *NewYork: Wiley,* 1**,** 978.

DILLON, W. R., KUMAR, A. & MULANI, N. 1987. Offending estimates in covariance structure analysis: Comments on the causes of and solutions to Heywood cases. *Psychological Bulletin,* 101**,** 126.

DONALD, B. 1987. Multiple imputation for nonresponse in surveys. New York: J. Wiley & Sons.

DONG, Y. & PENG, C.-Y. J. 2013. Principled missing data methods for researchers. *SpringerPlus,* 2**,** 222.

DONNELLY, L. R. & MACENTEE, M. I. 2012. Social interactions, body image and oral health among institutionalised frail elders: an unexplored relationship. *Gerodontology,* 29**,** e28-e33.

DOUGLAS DE OLIVEIRA, D. W., MARQUES, D. P., AGUIAR-CANTUÁRIA, I. C., FLECHA, O. D. & GONÇALVES, P. F. 2013. Effect of surgical defect coverage on cervical dentin hypersensitivity and quality of life. *Journal of Periodontology,* 84**,** 768-775.

DUPONT, W. D. 2014. Statistical Modeling for Biomedical Researchers. *2nd Edition, Cambridge University Press.*

DURHAM, J., STEELE, J., WASSELL, R., EXLEY, C., MEECHAN, J., ALLEN, P. & MOUFTI, M. 2011. Creating a patient-based condition-specific outcome measure for Temporomandibular Disorders (TMDs): Oral Health Impact Profile for TMDs (OHIP-TMDs). *Journal of Oral Rehabilitation,* 38**,** 871-883.

EISEN, S. V., RANGANATHAN, G., SEAL, P. & SPIRO III, A. 2007. Measuring clinically meaningful change following mental health treatment. *The Journal of Behavioral Health Services & Research,* 34**,** 272-289.

EKANAYAKE, L. & PERERA, I. 2003. Validation of a Sinhalese translation of the Oral Health Impact Profile-14 for use with older adults. *Gerodontology,* 20**,** 95-99.

ELLIS, J., THOMASON, J., JEPSON, N., NOHL, F., SMITH, D. & ALLEN, P. 2008. A randomized-controlled trial of food choices made by edentulous adults. *Clinical Oral Implants Research,* 19**,** 356-361.

ENOKI, K., IKEBE, K., MATSUDA, K., YOSHIDA, M., MAEDA, Y. & THOMSON, W. 2013. Determinants of change in oral health-related quality of life over 7 years among older Japanese. *Journal of Oral Rehabilitation,* 40**,** 252-257.

FELCE, D. & PERRY, J. 1995. Quality of life: Its definition and measurement. *Research in Developmental Disabilities,* 16**,** 51-74.

FERGUSON, R. J., ROBINSON, A. B. & SPLAINE, M. 2002. Use of the reliable change index to evaluate clinical significance in SF-36 outcomes. *Quality of Life Research,* 11**,** 509-516.

FLØLO, T. N., ANDERSEN, J. R., NIELSEN, H. J. & NATVIG, G. K. 2014. Translation, adaptation, validation and performance of the American Weight Efficacy Lifestyle Questionnaire Short Form (WEL-SF) to a Norwegian version: a cross-sectional study. *PeerJ,* 2**,** e565.

FRUCHTER, B. 1954. Introduction to factor analysis. *D. Van Nostrand Co.*

FUEKI, K., IGARASHI, Y., MAEDA, Y., BABA, K., KOYANO, K., SASAKI, K., AKAGAWA, Y., KUBOKI, T., KASUGAI, S. & GARRETT, N. 2015. Effect of prosthetic restoration on oral health-related quality of life in patients with shortened dental arches: a multicentre study. *Journal of Oral Rehabilitation,* 42**,** 701-708.

GAGLIARDI, D., SLADE, G. & SANDERS, A. 2008. Impact of dental care on oral health-related quality of life and treatment goals among elderly adults. *Australian Dental Journal,* 53**,** 26-33.

GANDEK, B., SINCLAIR, S. J., KOSINSKI, M. & WARE JR, J. E. 2004. Psychometric evaluation of the SF-36® health survey in medicare managed care. *Health Care Financing Review,* 25**,** 5.

GARSON, D. 2008a. Factor Analysis: Statnotes. North Carolina State University Public Administration Program.

GARSON, D. G. 2008b. Correspondence analysis, from Statnotes: Topics in Multivariate Analysis. Retrieved 02/08/2011 from http://faculty.chass.ncsu.edu/garson/pa765/statnote.htm

GIAVARINA, D. 2015. Understanding bland altman analysis. *Biochemia Medica,* 25**,** 141-151.

GIFT, H. C. & ATCHISON, K. A. 1995. Oral health, health, and health-related quality of life. *Medical Care***,** NS57-NS77.

GILBERT, G. H. 2005. Racial and socioeconomic disparities in health from population-based research to practice-based research: the example of oral health. *Journal of Dental Education,* 69**,** 1003-1014.

GILBERT, G. H., DUNCAN, R. P., HEFT, M. W., DOLAN, T. A. & VOGEL, W. B. 1998a. Multidimensionality of oral health in dentate adults. *Medical Care,* 36**,** 988-1001.

GILBERT, G. H., DUNCAN, R. P. & VOGEL, W. B. 1998b. Determinants of dental care use in dentate adults: six-monthly use during a 24-month period in the Florida Dental Care Study. *Social Science & Medicine,* 47**,** 727-737.

GISLER, V., BASSETTI, R., MERICSKE-STERN, R., BAYER, S. & ENKLING, N. 2012. A cross-sectional analysis of the prevalence of dental anxiety and its relation to the oral health-related quality of life in patients with dental treatment needs at a university clinic in Switzerland. *Gerodontology,* 29**,** e290-e296.

GOMES, M., GUTACKER, N., BOJKE, C. & STREET, A. 2016. Addressing Missing Data in Patient-Reported Outcome Measures (PROMS): Implications for the Use of PROMS for Comparing Provider Performance. *Health Economics,* 25**,** 515-528.

GOOCH, B., DOLAN, T. & BOURQUE, L. 1989. Correlates of self-reported dental health status upon enrollment in the Rand Health Insurance Experiment. *Journal of Dental Education,* 53**,** 629-637.

GORSUCH, R. 1983. Factor analysis, 2nd. *Hillsdale, NJ: LEA*.

GRAHAM, J. W. 2009. Missing data analysis: Making it work in the real world. *Annual Review of Psychology,* 60**,** 549-576.

GRAHAM, J. W., OLCHOWSKI, A. E. & GILREATH, T. D. 2007. How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prevention Science,* 8**,** 206-213.

GREGORY, J., GIBSON, B. & ROBINSON, P. G. 2005. Variation and change in the meaning of oral health related quality of life: a 'grounded'systems approach. *Social Science & Medicine,* 60**,** 1859-1868.

GUARNIZO-HERREÑO, C. C., WATT, R. G., FULLER, E., STEELE, J. G., SHEN, J., MORRIS, S., WILDMAN, J. & TSAKOS, G. 2014. Socioeconomic position and subjective oral health: findings for the adult population in England, Wales and Northern Ireland. *BMC Public Health,* 14**,** 1.

GUILFORD, J. P. 1954. *Psychometric methods*, McGraw-Hill.

GUPTA, E. 2004. The determinants of Oral Health Related Quality of Life in Adults. *Phd Thesis. University of Sheffield.*

GUPTA, E., ROBINSON, P., MARYA, C. & BAKER, S. 2015. Oral Health Inequalities Relationships between Environmental and Individual Factors. *Journal of Dental Research,* 94**,** 1362-1368.

GUYATT, G., WALTER, S. & NORMAN, G. 1987. Measuring change over time: assessing the usefulness of evaluative instruments. *Journal of Chronic Diseases,* 40**,** 171-178.

GUYATT, G. H., FEENY, D. H. & PATRICK, D. L. 1993. Measuring health-related quality of life. *Annals of Internal Medicine,* 118**,** 622-629.

GUYATT, G. H., FERRANS, C. E., HALYARD, M. Y., REVICKI, D. A., SYMONDS, T. L., VARRICCHIO, C. G., KOTZEVA, A., VALDERAS, J. M., ALONSO, J. L. & GROUP, C. S. C. M. Exploration of the value of health-related quality-of-life information from clinical research and into clinical practice. Mayo Clinic Proceedings, 2007. Elsevier, 1229-1239.

GUZELDEMIR, E., TOYGAR, H. U., TASDELEN, B. & TORUN, D. 2009. Oral health–related quality of life and periodontal health status in patients undergoing hemodialysis. *The Journal of the American Dental Association,* 140**,** 1283-1293.

HAGEMAN, W. J. & ARRINDELL, W. A. 1993. A further refinement of the reliable change (RC) index by improving the pre-post difference score: Introducing RC ID. *Behaviour Research and Therapy,* 31**,** 693-700.

HAHN, E. A. 2011. Multi-item Scales and Tests: Development and Validation Methods. [Available from : http://www.feinberg.northwestern.edu/sites/bcc/docs/2011-lectures/Lecture5hahnmultiitemscales.pdf. Accessed on 25th April 2017].

HARDT, J., HERKE, M. & LEONHART, R. 2012. Auxiliary variables in multiple imputation in regression with missing X: a warning against including too many in small sample research. *BMC Medical Research Methodology,* 12**,** 184.

HASSAN, W. N. W., YUSOF, Z. Y. M., SHAHIDAN, S. S. Z., ALI, S. F. M. & MAKHBUL, M. Z. M. 2017. Validation and reliability of the translated Malay version of the psychosocial impact of dental aesthetics questionnaire for adolescents. *Health and Quality of Life outcomes,* 15**,** 23.

HASSEL, A., STEUKER, B., ROLKO, C., KELLER, L., RAMMELSBERG, P. & NITSCHKE, I. 2010. Oral health-related quality of life of elderly Germans--comparison of GOHAI and OHIP-14. *Community Dental Health,* 27**,** 242-247.

HATCHER, L. 1994. *A Step-by-Step Approach to Using the SAS System for Factor Analysis and Structural Equation Modeling,* SAS Publishing.

HELKIMO, M. 1974. Studies on function and dysfunction of the masticatory system. II. Index for anamnestic and clinical dysfunction and occlusal state. *Svensk tandläkare tidskrift. Swedish Dental Journal,* 67**,** 101.

HONGXING, L., LIST, T., NILSSON, M., JOHANSSON, A. & ASTRØM, A. N. 2014. Validity and reliability of OIDP and OHIP-14: a survey of Chinese high school students. *BMC Oral Health,* 14**,** 158.

HORN, J. L. 1967. On subjectivity in factor analysis. *Educational and Psychological Measurement,* 27**,** 811-820.

HORN, J. L. & KNAPP, J. R. 1974. Thirty wrongs do not make a right: Reply to Guilford. *Psychological Bulletin,* 81**,** 502-504.

HORN, J. L. & MCARDLE, J. J. 1980. Perspectives on Mathematical/Statistical Model Building (MASMOB) in research on aging. *In L. W. Poon (Ed.), Aging in the 1980s: Psychological Issues* 503-541.

HULME, C., ROBINSON, P., DOUGLAS, G., BAXTER, P., GIBSON, B., GODSON, J., VINALL-COLLIER, K., SALONIKI, E., MEADS, D. & BRUNTON, P. 2016. The INCENTIVE Study: a mixed methods evaluation of an innovation in commissioning and delivery of primary dental care compared to traditional dental contracting. *Health Services and Delivery Research,* 4.

HUMPHREYS, L. G., ILGEN, D., MCGRATH, D. & MONTANELLI, R. 1969. Capitalization on chance in rotation of factors. *Educational and Psychological Measurement,* 29**,** 259-271.

HUTCHESON, G. D. & SOFRONIOU, N. 1999. *The multivariate social scientist: Introductory statistics using generalized linear models,* Sage Publications.

HYLAND, M. E. 2003. A brief guide to the selection of quality of life instrument. *Health and Quality of Life Outcomes,* 1**,** 24.

IKEBE, K., HAZEYAMA, T., ENOKI, K., MURAI, S., OKADA, T., KAGAWA, R., MATSUDA, K.-I. & MAEDA, Y. 2012. Comparison of GOHAI and OHIP-14 measures in relation to objective values of oral function in elderly Japanese. *Community Dentistry and Oral Epidemiology,* 40**,** 406-414.

IKEBE, K., MATSUDA, K.-I., MORII, K., WADA, M., HAZEYAMA, T., NOKUBI, T. & ETTINGER, R. L. 2007. Impact of dry mouth and hyposalivation on oral health-related quality of life of elderly Japanese. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology,* 103**,** 216-222.

INGLEHART, M. R. & BAGRAMIAN, R. 2002. *Oral health-related quality of life,* Quintessence Publications.

JACOBSON, N. S., FOLLETTE, W. C. & REVENSTORF, D. 1984. Psychotherapy outcome research: Methods for reporting variability and evaluating clinical significance. *Behavior Therapy,* 15**,** 336-352.

JAIN, M., KAIRA, L., SIKKA, G., SINGH, S., GUPTA, A., SHARMA, R., SAWLA, L. & MATHUR, A. 2012. How do age and tooth loss affect oral health impacts and quality of life? A study comparing two state samples of gujarat and rajasthan. *Journal of Dentistry of Tehran University of Medical Sciences,* 9**,** 135-44.

JENKINSON, C., GRAY, A., DOLL, H., LAWRENCE, K., KEOGHANE, S. & LAYTE, R. 1997. Evaluation of index and profile measures of health status in a randomized controlled trial: comparison of the Medical Outcomes Study 36-Item Short Form Health Survey, EuroQol, and disease specific measures. *Medical Care,* 35**,** 1109-1118.

JENKINSON, C., LAYTE, R., COULTER, A. & WRIGHT, L. 1996. Evidence for the sensitivity of the SF-36 health status measure to inequalities in health: results from the Oxford healthy lifestyles survey. *Journal of Epidemiology and Community Health,* 50**,** 377-380.

JOHN, M. T., FEUERSTAHLER, L., WALLER, N., BABA, K., LARSSON, P., ČELEBIĆ, A., KENDE, D., RENER-SITAR, K. & REIßMANN, D. R. 2014. Confirmatory factor analysis of the oral health impact profile. *Journal of Oral Rehabilitation,* 41**,** 644-652.

JOHN, M. T., KOEPSELL, T. D., HUJOEL, P., MIGLIORETTI, D. L., LERESCHE, L. & MICHEELIS, W. 2004a. Demographic factors, denture status and oral health-related quality of life. *Community Dentistry and Oral Epidemiology,* 32**,** 125-132.

JOHN, M. T., PATRICK, D. L. & SLADE, G. D. 2002. The German version of the Oral Health Impact Profile–translation and psychometric properties. *European Journal of Oral Sciences,* 110**,** 425-433.

JOHN, M. T., REIßMANN, D. R., SZENTPÉTERY, A. & STEELE, J. 2009. An approach to define clinical significance in prosthodontics. *Journal of Prosthodontics,* 18**,** 455-460.

JOHN, M. T., SLADE, G. D., SZENTPÉTERY, A. & SETZ, J. M. 2004b. Oral Health Related Quality of Life in Patients Treated with Fixed, Removable, and Complete Dentures 1 Month and 6 to 12 Months After Treatment. *International Journal of Prosthodontics,* 17.

JÖNSSON, B. & ÖHRN, K. 2014. Evaluation of the effect of non-surgical periodontal treatment on oral health-related quality of life: estimation of minimal important differences 1 year after treatment. *Journal of Clinical Periodontology,* 41**,** 275-282.

KAMPHAUS, R. W. & FRICK, P. J. 2005. *Clinical assessment of child and adolescent personality and behavior,* Springer Science & Business Media.

KAVIN, T., JAGADESAN, A. G. P. & VENKATARAMAN, S. S. 2012. Changes in quality of life and impact on patients' perception of esthetics after orthognathic surgery. *Journal of Pharmacy And Bioallied Sciences,* 4**,** 290.

KAZIS, L. E., ANDERSON, J. J. & MEENAN, R. F. 1989. Effect sizes for interpreting changes in health status. *Medical Care,* 27**,** S178-S189.

KEITH, T. Z. 1990. Confirmatory and hierarchical confirmatory analysis of the Differential Ability Scales. *Journal of Psychoeducational Assessment,* 8**,** 391-405.

KIEFFER, J., VAN WIJK, A., HO, J. & LINDEBOOM, J. 2012. The internal responsiveness of the Oral Health Impact Profile-14 to detect differences in clinical parameters related to surgical third molar removal. *Quality of Life Research,* 21**,** 1241-1247.

KIMURA, A., ARAKAWA, H., NODA, K., YAMAZAKI, S., HARA, E., MINO, T., MATSUKA, Y., MULLIGAN, R. & KUBOKI, T. 2012a. Response shift in oral health-related quality of life measurement in patients with partial edentulism. *Journal of Oral Rehabilitation,* 39**,** 44-54.

KIMURA, A., ARAKAWA, H., NODA, K., YAMAZAKI, S., HARA, E., MINO, T., MATSUKA, Y., MULLIGAN, R. & KUBOKI, T. 2012b. Response shift in oral health-related quality of life measurement in patients with partial edentulism. *Journal of oral rehabilitation,* 39**,** 44-54.

KLINE, P. 1979. *Psychometrics and Psychology,* Acad. Press.

KOLENIKOV, S. & BOLLEN, K. A. 2012. Testing Negative Error Variances Is a Heywood Case a Symptom of Misspecification? *Sociological Methods & Research,* 41**,** 124-167.

KRANJČIĆ, J., MIKUŠ, A., PERŠIĆ, S. & VOJVODIĆ, D. 2014. Factors Affecting Oral Health–Related Quality of Life Among Elderly Croatian Patients. *Acta Stomatologica Croatica,* 48**,** 174-182.

KRESSIN, N. 1996. Associations among different assessments of oral health outcomes. *Journal of Dental Education,* 60**,** 501-507.

KRESSIN, N., SPIRO III, A., BOSSÉ, R., GARCIA, R. & KAZIS, L. 1996. Assessing oral health-related quality of life: findings from the normative aging study. *Medical Care,* 34**,** 416-427.

KRISDAPONG, S. & SHEIHAM, A. 2014. Which aspects of an oral health–related quality of life measure are mainly associated with global ratings of oral health in children? *Community Dentistry and Oral Epidemiology,* 42**,** 129-138.

LAHTI, S., SUOMINEN-TAIPALE, L. & HAUSEN, H. 2008. Oral health impacts among adults in Finland: competing effects of age, number of teeth, and removable dentures. *European Journal of Oral Sciences,* 116**,** 260-266.

LARSSON, P., LIST, T., LUNDSTRÖM, I., MARCUSSON, A. & OHRBACH, R. 2004. Reliability and validity of a Swedish version of the Oral Health Impact Profile (OHIP-S). *Acta Odontologica Scandinavica,* 62**,** 147-152.

LAWLEY, D. & MAXWELL, A. 1971. Factor Analysis as a Statistical Methods (2nd ed.). *New York: American Elsevier Publishing Company, Inc., Publishers*.

LEAO, A. & SHEIHAM, A. 1996. The development of a socio-dental measure of dental impacts on daily living. *Community Dental Health,* 13**,** 22-26.

LEPLEGE, A. & HUNT, S. 1997. The problem of quality of life in medicine. *Journal of American Medical Association,* 278**,** 47-50.

LIANG, M. H. 2000. Longitudinal construct validity: establishment of clinical meaning in patient evaluative instruments. *Medical Care,* 38**,** II-84-II-90.

LITTLE, R. J. 1988. A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association,* 83**,** 1198-1202.

LITTLE, R. J. & RUBIN, D. B. 2014. *Statistical analysis with missing data*, John Wiley & Sons.

LIU, L.-J., XIAO, W., HE, Q.-B. & JIANG, W.-W. 2012. Generic and oral quality of life is affected by oral mucosal diseases. *BMC Oral Health,* 12**,** 1.

LO COCO, G., CHIAPPELLI, M., BENSI, L., GULLO, S., PRESTANO, C. & LAMBERT, M. J. 2008. The factorial structure of the outcome questionnaire-45: a study with an Italian sample. *Clinical Psychology & Psychotherapy,* 15**,** 418-423.

LOCKER, D. 1988. Measuring oral health: a conceptual framework. *Community Dental Health,* 5**,** 3-18.

LOCKER, D. 1992. The burden of oral disorders in a population of older adults. *Community Dental Health,* 9**,** 109-124.

LOCKER, D. 1997. Concepts of oral health, disease and the quality of life. *Measuring Oral Health and Quality of Life***,** 11-24.

LOCKER, D. 1998. Issues in measuring change in self-perceived oral health status. *Community Dentistry and Oral Epidemiology,* 26**,** 41-47.

LOCKER, D. 2003. Dental status, xerostomia and the oral health-related quality of life of an elderly institutionalized population. *Special Care in Dentistry,* 23**,** 86-93.

LOCKER, D. 2004. Oral health and quality of life. *Oral Health & Preventive Dentistry,* 2**,** 247-253.

LOCKER, D. & ALLEN, F. 2007a. What do measures of 'oral health-related quality of life' measure? *Community Dentistry and Oral Epidemiology,* 35**,** 401-411.

LOCKER, D. & ALLEN, F. 2007b. What do measures of 'oral health related quality of life'measure? *Community Dentistry and Oral Epidemiology,* 35**,** 401-411.

LOCKER, D. & ALLEN, P. F. 2002. Developing Short-form Measures of Oral Health-related Quality of Life. *Journal of Public Health Dentistry,* 62**,** 13-20.

LOCKER, D., JOKOVIC, A. & CLARKE, M. 2004. Assessing the responsiveness of measures of oral health related quality of life. *Community Dentistry and Oral Epidemiology,* 32**,** 10-18.

LOCKER, D., MATEAR, D., STEPHENS, M., LAWRENCE, H. & PAYNE, B. 2001. Comparison of the GOHAI and OHIP-14 as measures of the oral health-related quality of life of the elderly. *Community Dentistry and Oral Epidemiology,* 29**,** 373-381.

LOCKER, D. & MILLER, Y. 1994. Evaluation of subjective oral health status indicators. *Journal of Public Health Dentistry,* 54**,** 167-176.

LOCKER, D. & QUIÑONEZ, C. 2009. Functional and psychosocial impacts of oral disorders in Canadian adults: a national population survey. *Journal of the Canadian Dental Association,* 75.

LOCKER, D. & SLADE, G. 1993. Oral health and the quality of life among older adults: the oral health impact profile. *Journal of Canadian Dental Association,* 59**,** 830-3, 837-8, 844.

LOCKER, D. & SLADE, G. 1994. Association between clinical and subjective indicators of oral health status in an older adult population. *Gerodontology,* 11**,** 108-114.

LODI, G., CARROZZO, M., FURNESS, S. & THONGPRASOM, K. 2012. Interventions for treating oral lichen planus: a systematic review. *British Journal of Dermatology,* 166**,** 938-947.

LORD, F. M., NOVICK, M. R. & BIRNBAUM, A. 1968. Statistical theories of mental test scores. *Addison-Wesley*.

LUE, Y.-J., LIN, R.-F., CHEN, S.-S. & LU, Y.-M. 2009. Measurement of the functional status of patients with different types of muscular dystrophy. *The Kaohsiung Journal of Medical Sciences,* 25**,** 325-333.

MACCALLUM, R. C., BROWNE, M. W. & SUGAWARA, H. M. 1996. Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods,* 1**,** 130.

MACCALLUM, R. C., WIDAMAN, K. F., ZHANG, S. & HONG, S. 1999. Sample size in factor analysis. *Psychological Methods,* 4**,** 84.

MACCALLUM, R. C., ZHANG, S., PREACHER, K. J. & RUCKER, D. D. 2002. On the practice of dichotomization of quantitative variables. *Psychological Methods,* 7**,** 19.

MACEDO, C. G. D. & QUELUZ, D. D. P. 2011. Quality of life and self-perceived oral health among workers from a furniture industry. *Brazilian Journal of Oral Sciences,* 10**,** 226-232.

MAINDAL, H. T., SOKOLOWSKI, I. & VEDSTED, P. 2012. Adaptation, data quality and confirmatory factor analysis of the Danish version of the PACIC questionnaire. *The European Journal of Public Health,* 22**,** 31-36.

MCBEE, M. 2010. Modeling outcomes with floor or ceiling effects: An introduction to the Tobit model. *Gifted Child Quarterly,* 54**,** 314-320.

MCGRATH, C. & BEDI, R. 2002. Population based norming of the UK oral health related quality of life measure (OHQoL-UK©). *British Dental Journal,* 193**,** 521-524.

MCGRATH, C., WONG, A., LO, E. & CHEUNG, C. 2005. The sensitivity and responsiveness of an oral health related quality of life measure to tooth whitening. *Journal of Dentistry,* 33**,** 697-702.

MCHORNEY, C. A. 1999. Health status assessment methods for adults: past accomplishments and future challenges. *Annual Review of Public Health,* 20**,** 309-335.

MEHTA, A. & KAUR, G. 2011. Oral health-related quality of life—the concept, its assessment and relevance in dental research and education. *Indian Journal of Dentistry,* 2**,** 26-29.

MENARD, S. 1991. Longitudinal research: Quantitative applications in the social sciences. Newbury Park, CA: Sage.

MONTERO-MARTÍN, J., BRAVO-PÉREZ, M., ALBALADEJO-MARTÍNEZ, A., HERNÁNDEZ-MARTÍN, L. A. & ROSEL-GALLARDO, E. M. 2009. Validation the oral health impact profile (OHIP-14sp) for adults in Spain. *Med Oral Patol Oral Cir Bucal,* 14**,** E44-50.

MONTERO, J., BRAVO, M., VICENTE, M.-P., GALINDO, M.-P., LÓPEZ, J. F. & ALBALADEJO, A. 2010. Dimensional structure of the oral health-related quality of life in healthy Spanish workers. *Health and Quality of Life Outcomes,* 8**,** 24.

MOTALLEBNEJAD, M., HADIAN, H., MEHDIZADEH, S. & HAJIAHMADI, M. 2011. Validity and reliability of the Persian version of the oral health impact profile (OHIP)-14. *Caspian Journal of Internal Medicine,* 2**,** 314-320.

MUTHÉN, B. O. 1989. Tobit factor analysis†. *British Journal of Mathematical and Statistical Psychology,* 42**,** 241-250.

NAITO, M., YUASA, H., NOMURA, Y., NAKAYAMA, T., HAMAJIMA, N. & HANADA, N. 2006. Oral health status and health-related quality of life: a systematic review. *Journal of Oral Science,* 48**,** 1-7.

NG, S. K. & LEUNG, W. K. 2006. Oral health-related quality of life and periodontal status. *Community Dentistry and Oral Epidemiology,* 34**,** 114-122.

NIESTEN, D., WITTER, D., BRONKHORST, E. & CREUGERS, N. 2016. Validation of a Dutch version of the Geriatric Oral Health Assessment Index (GOHAI-NL) in care-dependent and care-independent older people. *BMC Geriatrics,* 16**,** 53.

NORRIS, M. & LECAVALIER, L. 2010. Evaluating the use of exploratory factor analysis in developmental disability psychological research. *Journal of Autism and Developmental Disorders,* 40**,** 8-20.

NORUSIS, M. 2005. SPSS 13.0 advanced statistical procedures companion. Chicago: Pretentice Hall. Inc.

NUTTALL, N., STEELE, J., PINE, C., WHITE, D. & PITTS, N. 2001. Adult dental health survey: The impact of oral health on people in the UK in 1998. *British Dental Journal,* 190**,** 121-126.

O'SULLIVAN, I., LADER, D., BEAVAN-SEYMOUR, C., CHENERY, V., FULLER, E. & SADLER, K. 2011. Foundation report: adult dental health survey 2009 (technical information). *Adult Dental Health Survey 2009*.

ÖHRN, K. & JÖNSSON, B. 2012. A comparison of two questionnaires measuring oral health-related quality of life before and after dental hygiene treatment in patients with periodontal disease. *International Journal of Dental Hygiene,* 10**,** 9-14.

OLIVEIRA, B. H. & NADANOVSKY, P. 2005. Psychometric properties of the Brazilian version of the Oral Health Impact Profile–short form. *Community Dentistry and Oral Epidemiology,* 33**,** 307-314.

ONG, A. D. & VAN DULMEN, M. H. 2007. Oxford handbook of methods in positive psychology. *Oxford University Press*.

OORD, E. J. & ARK, L. A. 1997. A note on the use of the Tobit approach for tests scores with floor or ceiling effects. *British Journal of Mathematical and Statistical Psychology,* 50**,** 351-364.

OORT, F. J. 2005. Using structural equation modeling to detect response shifts and true change. *Quality of Life Research,* 14**,** 587-598.

OSBORNE, J. W. 2012. *Best practices in data cleaning: A complete guide to everything you need to do before and after collecting your data*, Sage Publications.

PADIERNA, A., QUINTANA, J., AROSTEGUI, I., GONZALEZ, N. & HORCAJO, M. 2002. Changes in health related quality of life among patients treated for eating disorders. *Quality of Life Research,* 11**,** 545-552.

PASSALACQUA, A., REEVES, A., NEWTON, T., HUGHES, R., DUNNE, S., DONALDSON, N. & WILSON, N. 2012. An assessment of oral health promotion programmes in the United Kingdom. *European Journal of Dental Education,* 16**,** e19-e26.

PATRICK, D. L. & DEYO, R. A. 1989. Generic and disease-specific measures in assessing health status and quality of life. *Medical Care***,** S217-S232.

PATTUSSI, M. P., OLINTO, M. T. A., HARDY, R. & SHEIHAM, A. 2007. Clinical, social and psychosocial factors associated with self-rated oral health in Brazilian adolescents. *Community Dentistry and Oral Epidemiology,* 35**,** 377-386.

PERERA, I. & EKANAYAKE, L. 2008. Factors influencing perception of oral health among adolescents in Sri Lanka. *International Dental Journal,* 58**,** 349-355.

PETERSEN, P. E. 2003. The World Oral Health Report 2003: continuous improvement of oral health in the 21st century–the approach of the WHO Global Oral Health Programme. *Community Dentistry and Oral Epidemiology,* 31**,** 3-24.

PEUGH, J. L. & ENDERS, C. K. 2004. Missing data in educational research: A review of reporting practices and suggestions for improvement. *Review of Educational Research,* 74**,** 525-556.

PEYRE, H., LEPLÈGE, A. & COSTE, J. 2011. Missing data methods for dealing with missing items in quality of life questionnaires. A comparison by simulation of personal mean score, full information maximum likelihood, multiple imputation, and hot deck techniques applied to the SF-36 in the French 2003 decennial health survey. *Quality of Life Research,* 20**,** 287-300.

PIGOTT, T. D. 2001. A review of methods for missing data. *Educational Research and Evaluation,* 7**,** 353-383.

PIOVESAN, C., ÁBELLA, C. & ARDENGHI, T. M. 2011. Child oral health-related quality of life and socioeconomic factors associated with traumatic dental injuries in schoolchildren. *Oral Health & Preventive Dentistry,* 9.

PIOVESAN, C., BATISTA, A., FERREIRA, F. V. & ARDENGHI, T. M. 2009. Oral health-related quality of life in children: Conceptual issues. *Revista Odonto Ciência,* 24**,** 81-85.

QUEIROZ HERKRATH, A. P. C. D., HERKRATH, F. J., REBELO, M. A. B. & VETTORE, M. V. 2015. Measurement of health-related and oral health-related quality of life among individuals with nonsyndromic orofacial clefts: a systematic review and meta-analysis. *The Cleft Palate-Craniofacial Journal,* 52**,** 157-172.

QUINN, B. C., CATALANO, R. A. & FELBER, E. 2009. The Effect of Community-Level Unemployment on Preventive Oral Health Care Utilization. *Health Services Research,* 44**,** 162-181.

RABE-HESKETH, S. & SKRONDAL, A. 2012. Multilevel and Longitudinal Modeling Using Stata. *Stata Press books*.

REISSMANN, D. R., DIETZE, B., VOGELER, M., SCHMELZEISEN, R. & HEYDECKE, G. 2013. Impact of donor site for bone graft harvesting for dental implants on health-related and oral health-related quality of life. *Clinical Oral Implants Research,* 24**,** 698-705.

REVICKI, D., HAYS, R. D., CELLA, D. & SLOAN, J. 2008. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *Journal of Clinical Epidemiology,* 61**,** 102-109.

RIGDON, E. E. 1994. Calculating degrees of freedom for a structural equation model. *Structural Equation Modeling: A Multidisciplinary Journal,* 1**,** 274-278.

RINDSKOPF, D. 1984. Structural equation models empirical identification, Heywood cases, and related problems. *Sociological Methods & Research,* 13**,** 109-119.

ROBINSON, P. G., GIBSON, B., KHAN, F. A. & BIRNBAUM, W. 2003. Validity of two oral health-related quality of life measures. *Community Dentistry and Oral Epidemiology,* 31**,** 90-99.

ROWAN, K. 1994. Global questions and scores. *Measuring Health and Medical Outcomes***,** 54-76.

ROYSTON, P. 2004. Multiple imputation of missing values. *Stata Journal,* 4**,** 227-41.

RUBIN, D. 1987. Multivariate Imputation for Nonresponse in Surveys. New York: J. Wiley & Sons.

RUSANEN, J., LAHTI, S., TOLVANEN, M. & PIRTTINIEMI, P. 2009. Quality of life in patients with severe malocclusion before treatment. *The European Journal of Orthodontics,* 32**,** 43-48.

SANDERS, A. E., SLADE, G. D., LIM, S. & REISINE, S. T. 2009. Impact of oral disease on quality of life in the US and Australian populations. *Community Dentistry and Oral Epidemiology,* 37**,** 171-181.

SANTOS, C. M. D., OLIVEIRA, B. H. D., NADANOVSKY, P., HILGERT, J. B., CELESTE, R. K. & HUGO, F. N. 2013. The Oral Health Impact Profile-14:: a unidimensional scale? *Cadernos de Saude Publica,* 29**,** 749-757.

SATO, M. 1987. Pragmatic treatment of improper solutions in factor analysis. *Annals of the Institute of Statistical Mathematics,* 39**,** 443-455.

SAUB, R., LOCKER, D. & ALLISON, P. 2005. Derivation and validation of the short version of the Malaysian Oral Health Impact Profile. *Community Dentistry and Oral Epidemiology,* 33**,** 378-383.

SCHAFER, J. L. 1997. *Analysis of incomplete multivariate data,* CRC press.

SCHAFER, J. L. 1999. Multiple imputation: a primer. *Statistical Methods in Medical Research,* 8**,** 3-15.

SCULLY, C. 2000. Oral health in America: a report of the surgeon general. *Department of Health and Human Services, US Public Health Service*.

SEGÙ, M., COLLESANO, V., LOBBIA, S. & REZZANI, C. 2005. Cross-cultural Validation of a short form of the Oral Health Impact Profile for temporomandibular disorders. *Community Dentistry and Oral Epidemiology,* 33**,** 125-130.

SELIM, A. J., ROGERS, W., FLEISHMAN, J. A., QIAN, S. X., FINCKE, B. G., ROTHENDLER, J. A. & KAZIS, L. E. 2009. Updated US population standard for the Veterans RAND 12-item Health Survey (VR-12). *Quality of Life Research,* 18**,** 43-52.

SHORT, R. & HORN, J. 1984. Some notes on factor analysis of behavioral data. *Behaviour,* 90**,** 203-214.

SISCHO, L. & BRODER, H. 2011. Oral Health-related Quality of Life What, Why, How, and Future Implications. *Journal of Dental Research,* 90**,** 1264-1270.

SKARET, E., ASTROM, A. & HAUGEJORDEN, O. 2004. Oral Health Related Quality of Life (OHRQoL) review of existing instruments and suggestions for use in oral health outcome research in Europe. *European Global Health Indicators Development Project. Quintessence, Paris***,** 99-110.

SLADE, G., NUTTALL, N., SANDERS, A., STEELE, J., ALLEN, P. & LAHTI, S. 2005. Impacts of oral disorders in the United Kingdom and Australia. *British Dental Journal,* 198**,** 489.

SLADE, G. D. 1997. Derivation and validation of a short-form oral health impact profile. *Community Dentistry and Oral Epidemiology,* 25**,** 284-290.

SLADE, G. D. 1998. Assessing change in quality of life using the Oral Health Impact Profile. *Community Dentistry and Oral Epidemiology,* 26**,** 52-61.

SLADE, G. D. 2002. Assessment of oral health related quality of life. *In: Oral health-related quality of life. Inglehart MR, Bagramian RA, editors. Carol Stream, IL: Quintessence Publishing Co., Inc., pp. 29-46.*

SLADE, G. D. 2012. Oral health-related quality of life is important for patients, but what about populations? *Community Dentistry and Oral Epidemiology,* 40**,** 39-43.

SLADE, G. D. & SANDERS, A. E. 2011. The paradox of better subjective oral health in older age. *Journal of Dental Research,* 90**,** 1279-1285.

SLADE, G. D. & SPENCER, A. J. 1994. Development and evaluation of the Oral Health Impact Profile. *Community Dental Health,* 11**,** 3-11.

SOÁREZ, D., COELHO, P., CASTELO, A., ABRÃO, P., HOLMES, W. C. & CICONELLI, R. M. 2009. Brazilian-Portuguese translation and validation of the HIV/AIDS-Targeted Quality of Life Instrument. *Revista Panamericana de Salud Pública,* 25**,** 69-76.

STEELE, J. G., SANDERS, A. E., SLADE, G. D., ALLEN, P. F., LAHTI, S., NUTTALL, N. & SPENCER, A. J. 2004. How do age and tooth loss affect oral health impacts and quality of life? A study comparing two national samples. *Community Dentistry and Oral Epidemiology,* 32**,** 107-114.

STENMAN, U., AHLQWIST, M., BJÖRKELUND, C. & HAKEBERG, M. 2012. Oral health–related quality of life–associations with oral health and conditions in Swedish 70-year-old individuals. *Gerodontology,* 29**,** e440-e446.

STERNE, J. A., WHITE, I. R., CARLIN, J. B., SPRATT, M., ROYSTON, P., KENWARD, M. G., WOOD, A. M. & CARPENTER, J. R. 2009. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *British Medical Journal,* 338**,** b2393.

STONE, S. J., HEASMAN, P. A., STAINES, K. S. & MCCRACKEN, G. I. 2015. The impact of structured plaque control for patients with gingival manifestations of oral lichen planus: a randomized controlled study. *Journal of Clinical Periodontology,* 42**,** 356-362.

STRAUSS, R. P. & HUNT, R. J. 1993. Understanding the value of teeth to older adults: influences on the quality of life. *The Journal of the American Dental Association,* 124**,** 105-110.

STREINER, D. L. & NORMAN, G. R. 1989. *Health measurement scales: a practical guide to their development and use*, Oxford University Press, USA.

SUTINEN, S., LAHTI, S., NUTTALL, N. M., SANDERS, A. E., STEELE, J. G., ALLEN, P. F. & SLADE, G. D. 2007. Effect of a 1-month vs. a 12-month reference period on responses to the 14-item Oral Health Impact Profile. *European Journal of Oral Sciences,* 115**,** 246-249.

SZENTPÉTERY, A., SZABÓ, G., MARADA, G., SZÁNTÓ, I. & JOHN, M. T. 2006. The Hungarian version of the oral health impact profile. *European Journal of Oral Sciences,* 114**,** 197-203.

TABACHNICK, B. & FIDELL, L. 2013. Using Multivariate statistics, 6th edition. *Boston Pearson, - Always Learning*

TEIXEIRA, I. P., NOVAIS, I. D. P., PINTO, R. D. M. C. & CHEIK, N. C. 2012. Cultural adaptation and validation of the KINDL questionnaire in Brazil for adolescents between 12 and 16 years of age. *Revista Brasileira de Epidemiologia,* 15**,** 845-857.

TOBIN, J. 1958. Estimation of relationships for limited dependent variables. *Econometrica: Journal of the Econometric Society***,** 24-36.

TSAKOS, G., ALLEN, P. F., STEELE, J. G. & LOCKER, D. 2012. Interpreting oral health-related quality of life data. *Community Dentistry and Oral Epidemiology,* 40**,** 193-200.

TU, Y.-K., BLANCE, A., CLEREHUGH, V. & GILTHORPE, M. 2005. Statistical power for analyses of changes in randomized controlled trials. *Journal of Dental Research,* 84**,** 283-287.

VAN DRIEL, O. P. 1978. On various causes of improper solutions in maximum likelihood factor analysis. *Psychometrika,* 43**,** 225-243.

VANVOORHIS, C. W. & MORGAN, B. L. 2007. Understanding power and rules of thumb for determining sample sizes. *Tutorials in Quantitative Methods for Psychology,* 3**,** 43-50.

VIRDI, M. S. 2015. Emerging Trends in Oral Health Sciences and Dentistry. [Available from: https://www.intechopen.com/books/emerging-trends-in-oral-health-sciences-and-dentistry].

WANG, L., ZHANG, Z., MCARDLE, J. J. & SALTHOUSE, T. A. 2008. Investigating ceiling effects in longitudinal data analysis. *Multivariate Behavioral Research,* 43**,** 476-496.

WANG, T.-F., CHOU, C. & SHU, Y. 2013. Assessing the effects of oral health-related variables on quality of life in Taiwanese adults. *Quality of Life Research,* 22**,** 811-825.

WARE JR, J. E., BROOK, R. H., DAVIES, A. R. & LOHR, K. N. 1981. Choosing measures of health status for individuals in general populations. *American Journal of Public Health,* 71**,** 620-625.

WARE JR, J. E. & SHERBOURNE, C. D. 1992. The MOS 36-item short-form health survey (SF-36): I. Conceptual framework and item selection. *Medical Care***,** 473-483.

WENGER, N. K., MATTSON, M. E., FURBERG, C. D. & ELINSON, J. 1984. Assessment of quality of life in clinical trials of cardiovascular therapies. *The American Journal of Cardiology,* 54**,** 908-913.

WILLUMSEN, T., FJAERA, B. & EIDE, H. 2010. Oral health-related quality of life in patients receiving home-care nursing: associations with aspects of dental status and xerostomia. *Gerodontology,* 27**,** 251-257.

WILSON, I. B. & CLEARY, P. D. 1995. Linking clinical variables with health-related quality of life. *Journal of the American Medical Association,* 273**,** 59-65.

WOLF, E. J., HARRINGTON, K. M., CLARK, S. L. & MILLER, M. W. 2013. Sample size requirements for structural equation models an evaluation of power, bias, and solution propriety. *Educational and Psychological Measurement,* 73**,** 913-934.

WONG, A. H., CHEUNG, C. & MCGRATH, C. 2007. Developing a short form of Oral Health Impact Profile (OHIP) for dental aesthetics: OHIP- aesthetic. *Community Dentistry and Oral Epidemiology,* 35**,** 64-72.

WRIGHT, W. G., JONES, J. A., SPIRO III, A., RICH, S. E. & KRESSIN, N. R. 2009. Use of Patient Self-Report Oral Health Outcome Measures in Assessment of Dental Treatment Outcomes. *Journal of Public Health Dentistry,* 69**,** 95-103.

WYRWICH, K., NORQUIST, J., LENDERKING, W., ACASTER, S. & RESEARCH, I. A. C. O. I. S. F. Q. O. L. 2013. Methods for interpreting change over time in patient-reported outcome measures. *Quality of Life Research,* 22**,** 475-483.

YAO, D., JINSHENG, W., XIUBAI, Y., TIEWU, J., XIANHONG, W., KUN, Y., WEILONG, H., SHANWEN, O., SHIHAO, H. & XIAYU, C. 2010. Reliability and validity of SF-36 in advanced schistosomiasis. *Chinese Journal of Schistosomiasis Control,* 22**,** 40-46.

YONG, A. G. & PEARCE, S. 2013. A beginner's guide to factor analysis: Focusing on exploratory factor analysis. *Tutorials in Quantitative Methods for Psychology,* 9**,** 79-94.

YULE, P. L., DURHAM, J., PLAYFORD, H., MOUFTI, M. A., STEELE, J., STEEN, N., WASSELL, R. W. & OHRBACH, R. 2015. OHIP-TMDs: a patient-reported outcome measure for temporomandibular disorders. *Community Dentistry and Oral Epidemiology,* 43**,** 461-470.

ZHAO, N. 2009. The minimum sample size in factor analysis. *Retrieved March 23, 2017 from https://www.encorewiki.org/display/~nzhao/The+Minimum+Sample+Size+in+factor+Analysis*.

ZIEBLAND, S. 1994. Measuring changes in health status. *Measuring Health and Medical Outcomes***,** 42-53.

ZIĘTEK, M. & MALICKA, B. 2015. Validation of the Polish version of the oral health impact profile-14. *Health,* 3**,** 9-13.

ZIMMERMAN, D. W. & WILLIAMS, R. H. 1982. Gain scores in research can be highly reliable. *Journal of Educational Measurement,* 19**,** 149-154.

Appendices

Appendix A Patient Information Sheet

King's College Hospital NHS
NHS Foundation Trust

KCL Dental Institute
Bessemer Road
London SE5 9RS

Tel: 020 3299 9000
Fax: 020 3299 3185
www.kch.nhs.uk

INFORMATION SHEET FOR PARTICIPANTS

REC Reference Number:13/LO/0366

YOU WILL BE GIVEN A COPY OF THIS INFORMATION SHEET

Methodological Issues in Oral Health research - On the evaluation of Oral Health related Quality of Life

We would like to invite you to participate in this postgraduate (Ph.D) research. You should only participate if you want to; choosing not to take part will not disadvantage you in any way. Before you decide whether you want to take part, it is important for you to understand why the research is being done and what your participation will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information.

What are the aims of this research?

- We are measuring the oral health related quality of life, using a qualitative approach (where data are described in words) and a quantitative approach (data are collected in numbers) and we will be combine these two approaches to summarize the results.
- We will look how the oral health related quality of life will change over a time period. That is to check whether the quality of life increases or decreases or remains the same after some time.
- We will look how the new measure can be generalised to other situations and people.
- We will look how we can make a good substitution for data that are missing for different reasons. For example, when a patient does not want to answer a particular question.

Who is funding this research?

This research is funded by the Biostatistics & Research methods centre, Dental Institute, King's College London.

Who can participate in this study?

Patients attending King's College London Dental Institute, King's College Hospital Campus for dental treatment. They should be able to read, write and understand English, mentally sound, above 18 years of age. Also they should be able to give consent themselves without anybody's help.

What will you be asked to do?

If you agree to take part in this study, you will be asked to fill out a questionnaire about your dental health. For this, you will be given a project pack comprising information sheet, consent form and a questionnaire which consists of three parts: 1. Oral Health Impact Profile (14 items) 2. Attitude and Behaviour questions and 3. Demographic details.

When and where will the study take place?

Initially, you have to fill in the consent form and the questionnaire at the King's College London Dental Institute, King's College Hospital campus when you come for treatment. This may take 30

KCL Dental Institute
Bessemer Road
London SE5 9RS

Tel: 020 3299 9000
Fax: 020 3299 3185
www.kch.nhs.uk

minutes to complete. Then, the participants will be contacted by post or telephone and asked to fill in the same questionnaire as follow up after 2 and 4 months of the initial data collection.

Are there any risks involved in participating?

Since this study is questionnaire based, there are no potential risks involved for the patients and also the patients can withdraw from the study at any stage should they wish. However, patients personal information and postal information will be collected. The personal data will be handled according to UK Data protection act 1998. The personal information will not be exposed to any other persons except the researcher and the Supervisor(s).

Are there any benefits involved in participating?

This study will be helpful for the scientific community for the service improvement and for future research.

How will we maintain your privacy and confidentiality?

Once the questionnaire is completed, a dummy number will be given and any personal information will be safely kept and will not be available to anyone except the researcher and the supervisor(s). Data will be coded and stored in the computer which will be physically and electronically (password) protected.

Where will the results be published?

The results will be published as a Ph.D dissertation of the Researcher. Also, the results may be published in Scientific/ peer reviewed journals/ conferences. In both cases, only the summarised findings will be published. Individual data or any information that can lead to identification of the patient will not be disclosed.

What if I have questions about the project?

If you have questions about the project you can contact the researcher:
Mr A.Manoharan, Biostatistics & Res. Methods Centre, Floor 18, Tower wing, Guy's Hospital Campus, King's college London Dental Institute, London SE1 9RT.
Email: Manoharan.1.andiappan@kcl.ac.uk Phone: 020 7188 8091

It is up to you to decide whether to take part or not. If you decide to take part you are still free to withdraw from the study at any time and without giving a reason. A decision to withdraw at any time, or a decision not to take part, will not affect the standard of care you receive. If you do decide to take part you will be given this information sheet to keep and be asked to sign a consent form. Submission of a completed questionnaire implies consent to participate. As participation is anonymous it will not be possible for us to withdraw your data once you have returned your questionnaire.

If you have any questions or require more information about this study, please contact the researcher using the following contact details:

King's College Hospital
NHS Foundation Trust

KCL Dental Institute
Bessemer Road
London SE5 9RS

Tel: 020 3299 9000
Fax: 020 3299 3185
www.kch.nhs.uk

Manoharan Andiappan, Statistician, Biostatistics & Res. Methods centre, Floor 18, Tower Wing, Guy's Hospital campus, King's College London Dental Institute, London SE1 9RT. Phone: 020 7188 8091Email: manoharan.1.andiappan@kcl.ac.uk

If this study has harmed you in any way, you can contact King's College London using the details below for further advice and information:

Prof.Nora Donaldson, Biostatistics Unnit, Floor 18, Tower wing, Guy's Hospital. King's College London Dental Institute, London SE1 9RT. Phone: 020 7188 8091.

Local Contact:

Professor Stephen Dunne
Professor of Primary Dental Care
Dental Surgery Building Extension
King's College London Dental Institute
Denmark Hill
London SE5 9RW.

Thanks very much for your time. Once again please ask for more information on the project if you wish.

Professor Nora Donaldson, Principal Investigator.

Appendix B Patient Consent Form

King's College Hospital NHS
NHS Foundation Trust

Centre Number:

Study Number: 13/LO/0366

Patient Identification Number for this trial:

KCL Dental Institute
Bessemer Road
London SE5 9RS

Tel: 020 3299 9000
Fax: 020 3299 3185
www.kch.nhs.uk

CONSENT FORM

Methodological issues in oral health research - On the evaluation of the Oral Health Related Quality of Life

Name of the Researcher: Manoharan Andiappan

Please tick the box.

I confirm that I have read and understood the information sheet dated -14-03-2013 (version 2) for the above study and have had the opportunity to ask question. ☐

1. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my medical care or legal rights being affected. ☐
2. I understand that sections of any of my medical notes may be looked at by responsible individuals from King's College London or from regulatory authorities where it is relevant to me taking part in research. I give permission for these individuals to have access to my records. ☐
3. I understand that I will be contacted by post/ phone at 3 months and 6 months after the initial data collection for the follow up questionnaire. Also, I agree to give my postal address/ phone number/ email for the same. ☐
4. I agree to take part in the above study. ☐
5. I agree for the retention of my personal details, duly anonimized, for up to 3 years, for Phd research use. ☐
6. Any questions from me (the participant) have been answered in a satisfactory manner. ☐

Name of the Patient	Date	Signature
Name of the Person taking Consent	Date	Signature
Researcher	Date	Signature

Patient Consent form dated 07-03-2013. Version-2.0

INVESTOR IN PEOPLE

POSITIVE ABOUT DISABLED PEOPLE

Appendix C Questionnaire

survey 1

NHS

KING'S College LONDON

University of London

Dear Participant,
In accordance with ethical standards and regulations, all information will be treated with the strictest confidentiality and nothing that could reveal your identity will be available outside of the study centre. The clinicians will not see this information. It is my responsibility to answer any questions you may have. Please feel free to ask me any questions.

Thank you in advance for your important contribution to this research.

Kindest regards,

Manoharan Andiappan

Date the survey was completed: ______________________________

Date of upcoming treatment : ______________________________

Type of treatment: ☐ one tooth extraction ☐ one tooth filling ☐ root canal

Do you have dentures? ☐ No ☐ partial dentures ☐ full dentures

If you have dentures, do they feel loose in your mouth? ☐ Yes ☐ No

Unique participant code: ______________________________

version 2.0

Oral Health Impact Profile (14)

Please check one answer. Then check the level of importance it has to you.

In the past few days....

1a Have you had trouble pronouncing any words because of problems with your teeth, mouth or dentures?

☐ Never ☐ Hardly Ever ☐ Occasionally ☐ Fairly often ☐ Very often ☐ Always

1b If so, how much has it bothered you?

☐ Not at all ☐ Very little ☐ Little ☐ Much ☐ Very much ☐ Intolerable

---

2a Have you felt that your sense of taste has worsened because of problems with your teeth, mouth or dentures?

☐ Never ☐ Hardly Ever ☐ Occasionally ☐ Fairly often ☐ Very often ☐ Always

2b If so, how much has it bothered you?

☐ Not at all ☐ Very little ☐ Little ☐ Much ☐ Very much ☐ Intolerable

---

3a Have you had painful aching in your mouth?

☐ Never ☐ Hardly Ever ☐ Occasionally ☐ Fairly often ☐ Very often ☐ Always

3b If so, how much has it bothered you?

☐ Not at all ☐ Very little ☐ Little ☐ Much ☐ Very much ☐ Intolerable

---

4a Have you found it uncomfortable to eat any foods because of problems with your teeth, mouth or dentures?

☐ Never ☐ Hardly Ever ☐ Occasionally ☐ Fairly often ☐ Very often ☐ Always

4b If so, how much has it bothered you?

☐ Not at all ☐ Very little ☐ Little ☐ Much ☐ Very much ☐ Intolerable

---

5a Have you felt self-conscious because of problems with your teeth, mouth or dentures?

☐ Never ☐ Hardly Ever ☐ Occasionally ☐ Fairly often ☐ Very often ☐ Always

5b If so, how much has it bothered you?

☐ Not at all ☐ Very little ☐ Little ☐ Much ☐ Very much ☐ Intolerable

---

6a Have you felt tense because of problems with your teeth, mouth or dentures?

☐ Never ☐ Hardly Ever ☐ Occasionally ☐ Fairly often ☐ Very often ☐ Always

6b If so, how much has it bothered you?

☐ Not at all ☐ Very little ☐ Little ☐ Much ☐ Very much ☐ Intolerable

7a Has your diet been unsatisfactory because of problems with your teeth, mouth or dentures?

☐ Never ☐ Hardly Ever ☐ Occasionally ☐ Fairly often ☐ Very often ☐ Always

7b If so, how much has it bothered you?

☐ Not at all ☐ Very little ☐ Little ☐ Much ☐ Very much ☐ Intolerable

8a Have you had to interrupt meals because of problems with your teeth, mouth or dentures?

☐ Never ☐ Hardly Ever ☐ Occasionally ☐ Fairly often ☐ Very often ☐ Always

8b If so, how much has it bothered you?

☐ Not at all ☐ Very little ☐ Little ☐ Much ☐ Very much ☐ Intolerable

9a Have you found it difficult to relax because of problems with your teeth, mouth or dentures?

☐ Never ☐ Hardly Ever ☐ Occasionally ☐ Fairly often ☐ Very often ☐ Always

9b If so, how much has it bothered you?

☐ Not at all ☐ Very little ☐ Little ☐ Much ☐ Very much ☐ Intolerable

10a Have you been a bit embarrassed because of problems with your teeth, mouth or dentures?

☐ Never ☐ Hardly Ever ☐ Occasionally ☐ Fairly often ☐ Very often ☐ Always

10b If so, how much has it bothered you?

☐ Not at all ☐ Very little ☐ Little ☐ Much ☐ Very much ☐ Intolerable

11a Have you been a bit irritable with other people because of problems with your teeth, mouth or dentures?

☐ Never ☐ Hardly Ever ☐ Occasionally ☐ Fairly often ☐ Very often ☐ Always

11b If so, how much has it bothered you?

☐ Not at all ☐ Very little ☐ Little ☐ Much ☐ Very much ☐ Intolerable

12a Have you had difficulty doing your usual jobs because of problems with your teeth, mouth or dentures?

☐ Never ☐ Hardly Ever ☐ Occasionally ☐ Fairly often ☐ Very often ☐ Always

12b If so, how much has it bothered you?

☐ Not at all ☐ Very little ☐ Little ☐ Much ☐ Very much ☐ Intolerable

13a Have you felt that life in general was less satisfying because of problems with your teeth, mouth or dentures?

☐ Never ☐ Hardly Ever ☐ Occasionally ☐ Fairly often ☐ Very often ☐ Always

13b If so, how much has it bothered you?

☐ Not at all ☐ Very little ☐ Little ☐ Much ☐ Very much ☐ Intolerable

14a Have you been totally unable to function because of problems with your teeth, mouth or dentures?

☐ Never ☐ Hardly Ever ☐ Occasionally ☐ Fairly often ☐ Very often ☐ Always

14b If so, how much has it bothered you?

☐ Not at all ☐ Very little ☐ Little ☐ Much ☐ Very much ☐ Intolerable

Personal Details

I. Age : __________ years

II. Gender: ☐ Male ☐ Female

III. Ethnic Origin: (Please tick the one which most closely describes)

White
☐ British
☐ Irish
☐ Other

Mixed
☐ White & Black Caribbean
☐ White & Black African
☐ White & Asian
☐ Others

Asian or Asian British

☐ Indian
☐ Pakistan
☐ Bangladeshi
☐ Other

Black or Black British

☐ Caribbean
☐ African
☐ Others

Chinese or Other Ethnic group

☐ Chinese
☐ Other ethnic background (Specify) ___________________________

IV. Relationship status:

☐ Single ☐ In a relationship ☐ Married ☐ Separated

☐ Widowed ☐ Divorced

V. Education (Highest level achieved)

☐	GCSE	☐	A-Levels
☐	Degree (BSc, BA, etc)	☐	Advanced Degree (MSc, MA, etc)
☐	Research Degree	☐	Others (Specify)_________________

VI. Your last dental visit

☐	This is my first time	☐	Within last three months
☐	3-6 months before	☐	6-12 months before
☐	Before 1 year	☐	Don't remember

VII. Your profession

☐	Unemployed	☐	Part time
☐	Full time	☐	Student
☐	Others ------------------------ (Specify)		

VIII. Treatment Needs

☐	Periodontics	☐	Operatives
☐	Endodontics	☐	Surgery
☐	Orthodontics	☐	Prosthetics

If you do not want to answer any questions or likely to drop out from the study, please tick the reason(s) that apply.

IX. If you did not answer any of the questions or are likely to drop out of the study, please check any reasons that apply.

☐ I did not want to answer the questions. ☐ I was unable to answer the questions.
☐ Probably, I will not continue in this study.

Mailing Information

The purpose of the mailing information is to mail you the 2nd and 3rd version of this survey.

First and Last Name: ____

Mailing Address: ____

____

____

Is there a telephone number or email address
We can reach you at? ____________________________

This information is optional. The phone call is to confirm that our survey will arrive to you soon.

Best time to call you (Example: 6.30PM) ______________________________

The questionnaire is complete. Thank you for your time.

Appendix D Ethical Approval

NHS
Health Research Authority

NRES Committee London - Bromley
Bristol Research Ethics Committee Centre
Whitefriars
Level 3, Block B
Lewins Mead
Bristol
BS1 2NT

Telephone: 0117 342 1387
Facsimile: 0117 342 0445

06 March 2013

Professor Nora Donaldson
Professor of Biostatistics
King's College London
Biostatistics & Res Methods Centre,
King's College London Dental Institute,
Floor 18, Tower Wing, Guy's Hospital Campus, London.
SE1 9RT

Dear Professor Donaldson

Study title:	**Methodological Issues in Oral Health research - On the evaluation of Oral Health Related Quality of Life**
REC reference:	**13/LO/0366**
Protocol number:	**1**
IRAS project ID:	**118443**

The Proportionate Review Sub-Committee of the NRES Committee London - Bromley reviewed the above application on 04 March 2013.

Provisional opinion

The Sub-Committee would be content to give a favourable ethical opinion of the research, subject to the following changes being made to the documentation for study participants:

1) Changes to the Participant Information Sheet
a) Re-written in layman's language - particularly the Aims and Benefits section to make the information easier to read and user friendly

2) Changes to the Consent Form
a) Add a statement to agree for retention of personal details for up to 3 years for unspecified PHD use
b) Add a statement to indicate that any questions from the participant have been answered in a satisfactory manner

When submitting your response, please send the revised documentation underlining or otherwise highlighting the changes you have made and giving revised version numbers and dates.

Authority to consider your response and to confirm the final opinion on behalf of the Committee has been delegated to the Chair.

A Research Ethics Committee established by the Health Research Authority

Please let me know if you need any further clarification or would find it helpful to discuss the changes required with the lead reviewer.

The Committee will confirm the final ethical opinion within 7 days of receiving a full response.

Documents reviewed

The documents reviewed were:

Document	*Version*	*Date*
Covering Letter		21 February 2013
Evidence of insurance or indemnity		
Investigator CV		
Other: CV: Student		
Other: CV: second supervisor		
Other: CV: Third supervisor		
Participant Consent Form	1	14 December 2012
Participant Information Sheet	1	14 December 2012
Protocol	1	14 December 2012
Questionnaire	1.0	14 December 2012
REC application		

Membership of the Committee

The members of the Committee who were present at the meeting are listed on the attached sheet.

There were no declarations of interest

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

13/LO/0366	**Please quote this number on all correspondence**

Yours sincerely

pp

Ms Carol Jones
Chair

Email: nrescommittee.london-bromley@nhs.net

Enclosures: *List of names and professions of members who took part in the review*

Copy to: *Mr Keith Brennan*
Mr Zoe Harris, King's College Hospital NHS Foundation Trust

NRES Committee London - Bromley

Attendance at PRS Sub-Committee of the REC meeting on 04 March 2013

Committee Members:

Name	*Profession*	*Present*	*Notes*
Mr Munir Ahmed	Consultant Urologist	Yes	
Mrs Susan Beer	Retired Project Manager	Yes	
Ms Carol Jones (Chair)	Management Consultant	Yes	

Also in attendance:

Name	*Position (or reason for attending)*
Mrs Vicky Canfield-Duthie	Assistant Committee Coordinator

Appendix E   Ethical Approval (Amendment)

NHS
Health Research Authority
NRES Committee London - Bromley
Bristol Research Ethics Committee Centre
Whitefriars
Level 3, Block B
Lewins Mead
Bristol
BS1 2NT

Telephone: 0117 342 1387
Facsimile: 0117 342 0445

19 March 2013

Professor Nora Donaldson
Professor of Biostatistics
King's College London
Biostatistics & Res Methods Centre,
King's College London Dental Institute,
Floor 18, Tower Wing,
Guy's Hospital Campus,
London.
SE1 9RT

Dear Professor Donaldson

Study title:	**Methodological Issues in Oral Health research - On the evaluation of Oral Health Related Quality of Life**
REC reference:	**13/LO/0366**
Protocol number:	**1**
IRAS project ID:	**118443**

Thank you for your letter of 14th March 2013, responding to the Proportionate Review Sub-Committee's request for changes to the documentation for the above study.

The revised documentation has been reviewed and approved by the sub-committee.

We plan to publish your research summary wording for the above study on the NRES website, together with your contact details, unless you expressly withhold permission to do so. Publication will be no earlier than three months from the date of this favourable opinion letter. Should you wish to provide a substitute contact point, require further information, or wish to withhold permission to publish, please contact the Co-ordinator Mrs Vicky Canfield-Duthie, nrescommittee.london-bromley@nhs.net.

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation as revised.

Ethical review of research sites

The favourable opinion applies to all NHS sites taking part in the study, subject to management permission being obtained from the NHS/HSC R&D office prior to the start of the study (see "Conditions of the favourable opinion" below).

Conditions of the favourable opinion

The favourable opinion is subject to the following conditions being met prior to the start of the study.

Management permission or approval must be obtained from each host organisation prior to the start of the study at the site concerned.

Management permission ("R&D approval") should be sought from all NHS organisations involved in the study in accordance with NHS research governance arrangements.

Guidance on applying for NHS permission for research is available in the Integrated Research Application System or at http://www.rdforum.nhs.uk.

Where a NHS organisation's role in the study is limited to identifying and referring potential participants to research sites ("participant identification centre"), guidance should be sought from the R&D office on the information it requires to give permission for this activity.

For non-NHS sites, site management permission should be obtained in accordance with the procedures of the relevant host organisation.

Sponsors are not required to notify the Committee of approvals from host organisations.

It is the responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).

You should notify the REC in writing once all conditions have been met (except for site approvals from host organisations) and provide copies of any revised documentation with updated version numbers. The REC will acknowledge receipt and provide a final list of the approved documentation for the study, which can be made available to host organisations to facilitate their permission for the study. Failure to provide the final versions to the REC may cause delay in obtaining permissions.

Approved documents

The documents reviewed and approved by the Committee are:

Document	*Version*	*Date*
Covering Letter		21 February 2013
Evidence of insurance or indemnity		
Investigator CV		
Other: CV: Student		

Other: CV: second supervisor		
Other: CV: Third supervisor		
Participant Consent Form	2.0	07 March 2013
Participant Information Sheet	2.0	14 March 2013
Protocol	1	14 December 2012
Questionnaire	1.0	14 December 2012
REC application		
Response to Request for Further Information		14 March 2013

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

After ethical review

Reporting requirements

The attached document "After ethical review – guidance for researchers" gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Adding new sites and investigators
- Notification of serious breaches of the protocol
- Progress and safety reports
- Notifying the end of the study

The NRES website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

Feedback

You are invited to give your view of the service that you have received from the National Research Ethics Service and the application procedure. If you wish to make your views known please use the feedback form available on the website.

Further information is available at National Research Ethics Service website > After Review

13/LO/0366	**Please quote this number on all correspondence**

We are pleased to welcome researchers and R & D staff at our NRES committee members' training days – see details at http://www.hra.nhs.uk/hra-training/

With the Committee's best wishes for the success of this project.

Yours sincerely

pp V Confieldith

Ms Carol Jones
Chair

Email: nrescommittee.london-bromley@nhs.net

Enclosures: *"After ethical review – guidance for researchers" [SL-AR2]*

Copy to: *Mr Keith Brennan*

Mr Zoe Harris, King's College Hospital NHS Foundation Trust

Appendix F NHS R&D Approval

13 May 2013

King's College Hospital NHS
NHS Foundation Trust

Research and Development Department
First Floor Jennie Lee House, 34 Love Walk
Kings College Hospital NHS Trust
London SE5 8AD

Tel: 020 3299 9000
Fax: 020 3299 3445
Minicom: 020 3299 9009
www.kch.nhs.uk
kch-tr.research@nhs.net

Direct tel: 020 3299 1981
Direct fax: 020 3299 5515

King's College Hospital NHS
Dental Surgery Building Extension
Office 209, Caldecot Road
London
SW5 9RW

Dear Professor Donaldson,

Study Title: Methods to evaluate Oral Health Related Quality of Life

In accordance with the Department of Health's Research Governance Framework for Health and Social Care, all research projects taking place within the Trust must receive a favourable opinion from an ethics committee and approval from the Department of Research and Development (R&D) prior to commencement.

- **Ethics ref: 13/LO/0366**
- **Co-sponsors: KCL & KCH**
- **Funder: No cost No funding**
- **End date (as per IRAS application): 28/02/2014**
- **Protocol: Version 1 dated 14/12/2012**
- **Site: King's College NHS Foundation Trust**
- **R&D approval Date: 13 May 2013**

R&D have reviewed the documentation submitted for this project and I am pleased to inform you that we are approving the work to proceed within **King's College Hospital NHS Foundation Trust**. The study has been allocated the Trust R&D registration number **KCH13- 076.** Please quote this registration number in any communications with the R&D Department regarding your project.

Conditions of NHS Permission for research:

- The Principal Investigator must notify R&D of the actual end date of the project.
- The Principal Investigator is responsible for ensuring that Data Protection procedures are observed throughout the course of the project.
- The project must follow the agreed protocol and be conducted in accordance with all Trust Policies and Procedures especially those relating to research and data management.
- R&D must be notified of any changes to the protocol prior to implementation.
- Please submit a copy of the progress report on the anniversary of the Ethics favourable opinion **(19 March)**

If appropriate it is recommended that you register with the Current Controlled Trials website; http://isrctn.org/

Please ensure that you are aware of your responsibilities in relation to The Data Protection Act 1998, NHS Confidentiality Code of Practice, NHS Caldicott Report and Caldicott Guardians, the Human Tissue Act 2004, Good Clinical Practice, the NHS Research Governance Framework for Health and Social Care, Second Edition April 2005 and any further legislation released during the time of this study.
Members of the research team must have appropriate substantive or honorary contracts with the Trust prior to the study commencing. Any additional researchers who join the study at a later stage must also hold a suitable contract.

If the project is a clinical trial under the European Union Clinical Trials Directive the following must also be complied with:

1. The EU Directive on Clinical Trials (Directive 2001/20/EC) and UK's implementation of the Directive: The Medicines for Human Use (Clinical Trials) Regulations 2004;
2. The EU Directive on Principles and Guidelines for Good Clinical Practice (EU Commission Directive 2005/28/EC); and UK's implementation of the Directive: The Medicines for Human Use (Clinical Trials) Amendment Regulations 2006;

Amendments
Please ensure that you submit a copy of any amendments made to this study to the R&D Department.

Annual Report
It is obligatory that an annual report is submitted by the Chief Investigator to the research ethics committee, and we ask that a copy is sent to the R&D Department. The yearly period commences from the date of receiving a favourable opinion from the ethics committee.

Should you require any further information please do not hesitate to contact us.

In line with the Research Governance Framework, your project may be randomly selected for monitoring for compliance against the standards set out in the Framework. For information, the Trust's process for the monitoring of projects and the associated guidance is available from the Trust's intranet or on request from the R&D Department. You will be notified by the R&D Department if and when your project has been selected as part of the monitoring process. No action is needed until that time.

Many thanks for registering your research project

Yours sincerely,

Abdul Babalola
Research Governance Coordinator

cc. Sponsor, Keith Brennan, Room 1.8 Hodgkin Building, Guy's Campus, London. SE1 4UL

Appendix G Coding Scheme

All the OHIP -14 items were coded from 0 to 5 as below.

0 = Not at all

1= Very little

2 = Little

3 = Much

4 = Very much

5 = Intolerable

Oral Health Promotion Programme:

0 = No

1 = Yes

Personal Details:

Age – Actual value in years

Gender:

0 = Female

1 = Male

Ethnicity

1 = British

2 = Irish

3 = Other

4 = Indian

5 = Pakistani

6 = Bangladeshi

7 = Other

8= Chinese

9 = Other ethnic background

10 = White & Black Caribbean

11 = White & Black African

12 = White & Asian

13 = Other mixed

14 = Caribbean

15 = African

16 = Other Black or Black British

Relationship status

1 = Single

2 = In a relationship

3 = Married

4 = Separated

5 = Widowed

6 = Divorced

Education

1 = GCSE

2 = A levels

3 = Degree

4 = Advanced Degree

5 = Research Degree

6 = Others

Last Dental Visit

1 = First time

2 = Within last 3 months

3 = 3-6 months before

4 = 6-12 months before

5 = before 1 year

6 = Don’t remember

Profession

1 = Unemployed

2 = Full time

3 = Part time

4 =Student

5 = Retired

6 = Self employed

7 = Others

Treatment Needs

1 = Restoratives

(Periodontics, Endodontics, Prosthetics (root canal, bridging, Crown, Partial denture, tooth grinding, false teeth, filling, checking Gums and plates).

2. Orthodontics

3. Operatives/ Surgery (Operatives, Surgery, Implants and extraction)

4. General Check-up (Check-up, General treatment and Infection)

5. Multiple needs (Patients who required more than one of the above categories)

Appendix H Conferences / Meetings attended

Conference/ Meetings/ Seminar Presentations

This research project or part of it was presented in the following meetings/ seminars/ conferences.

1. Methodological issues in assessing Oral Health Related Quality of life using Oral Health Impact Profile (OHIP-14). Presented in King's College London 17th Annual Post Graduate Research Day held on 18th March 2015 at Floor 18, Dental Institute, GUYs Tower, London.

2. Methodological issues in assessing Oral Health Related Quality of life using Oral Health Impact Profile (OHIP-14). Presented at Patient and Population Health Divisional Post Graduate Research Day held on 7th July 2014, Denmark Hill campus, King's College London.

3. "Assessing Oral Health Related Quality of life and missing data using OHIP-14" by Manoharan Andiappan, Prof. Stephen Dunne and Prof. Nora Donaldson. A poster presented at ISOQOL 21st Annual conference: Quality of Life: Advancing measurement science and Transforming Healthcare held at Berlin, Germany from 15-18 October, 2014.

2\. Manoharan Andiappan and Nora Donaldson. "Relationship between the patients dental Visit and Oral Health Related Quality of life". A poster presentation at "93rd General Session & Exhibition of the International Association for Dental Research (IADR)" held in conjunction with the 44th Annual Meeting of the American Association for Dental Research (AADR) and the 39th Annual Meeting of the Canadian Association for Dental Research (CADR) at Boston, Mass., USA from March 11-14, 2015.

Courses Attended

1. Latent Variable Modelling of Categorical data – Tools of Analysis for Cross-National Surveys. – Training workshop held at London School of Economics and Political Science (Funded by Economic & Social research council). From 24/4/2012 to 25/4/2012.

2. Good Clinical Practice (online course). Completed on 06/06/2012.

3. Advanced Excel. One day course Conducted by King's College London on 14/6/2012.

3. Design and Analysis of Trials in Rare Cancers Joint Meeting of the Royal Statistical Society Medical Section and the MRC Hubs for Trials Methodology Research and the European Network for Cancer Research in Children and Adolescents and the International Rare Cancers Initiative. One day workshop on 05th October 2012.

4. "Showcasing Early Career Researchers in Statistical Computing" one day seminar organized by Royal Statistical Society, 18-10-2012.

5. Attended a course on "Research Design and Project Evaluation" (a course for PhD students) in Oct-Nov 2012 at King's College London.

6. Attended a course on "Reviewing the Literature" conducted by King's college London from Nov to Dec 2012.

7. A course on “Peer Review & How to write a Great Research Paper” conducted by Elsevier workshop - Scientific Journals held at King’s College London on 30th April 2014.

8. “Structural Equation Modelling for Cross sectional data” held at University of Southampton from 24/6/2014 to 25/6/2014.

9. Elsevier workshop: “Publish or Perish” and Journal Metrics – Implications for Academic careers. Conducted by King’s college London on 10-06-2015.

Appendix H Publication

1. Andiappan M, Hughes FJ, Dunne S, Gao W and Donaldson ANA. Adjusting the oral health related quality of life measure (using OHIP-14) for floor and ceiling effects. Journal of Oral Health and Community Dentistry. (2015). Vol 9(3): 99-104.